bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1 # Combining metabarcoding and morphological approaches to identify
2 # phytoplankton taxa associated with harmful algal blooms

3

4 Svetlana Esenkulova[1,2], Amy Tabata[1], Ben J.G. Sutherland[1], Nicola Haigh[3], Christopher M.
5 Pearce[1], Kristina M. Miller[1]
6 [1] Fisheries and Oceans Canada, Pacific Biological Station, Nanaimo, British Columbia, Canada
7 [2] Pacific Salmon Foundation, Vancouver, British Columbia, Canada
8 [3] Microthalassia Inc., Nanaimo, British Columbia, Canada
9
10 *Author for correspondence: SE
11 Pacific Salmon Foundation, #300 1682 West 7th Avenue, Vancouver, BC, Canada, V6J 4S6
12 Phone: 1 (250) 797-0705
13 Email: svesen@uvic.ca
14
15
16
17
18
19
20
21
22
23 **Data Deposition**: Raw sequence was uploaded to SRA under BioProject PRJNA544881 within
24 BioSample accessions SAMN11865982-SAMN11866125.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

25 # ABSTRACT

26 Impacts of harmful algal blooms (HABs) have increased in frequency, intensity, and
27 geographical distribution in the last several decades. Detection methods to date largely depend
28 on microscopic observations which require expertise and time-intensive processes. In this study,
29 we apply microscopic observational methods, quantitative real-time polymerase chain reaction
30 (qPCR), and metabarcoding with multiple markers (*i.e.* 16S, 18S-dinoflagellate and 18S-diatom,
31 and large subunit (LSU) rDNA) on cultured (N=30) and field (N=24) samples containing
32 suspected harmful algae (*e.g*., *Alexandrium* spp., *Chattonella* sp., *Chrysochromulina* spp.,
33 *Dictyocha* spp., *Heterosigma akashiwo*, *Protoceratium reticulatum*, *Pseudochattonella*
34 *verruculosa, Pseudo-nitzschia* spp., and *Pseudopedinella* sp). Good detectability was found
35 using previously published TaqMan assays for *A. tamarense*, *H. akashiwo*, and *P. verruculosa*.
36 Overall, the multiple marker metabarcoding results were superior to the morphology-based
37 method for detection and identification of harmful taxa, with the notable exception of taxa from
38 the silicoflagellate group (*e.g. Dictyocha* spp.), which had better detection by morphology.
39 Metabarcoding results depended greatly on the marker type applied, which highlights the value
40 of a multiple-marker approach. The combined results of the 18S and the LSU markers closely
41 corresponded with morphological identification of targeted species and provided the best overall
42 taxonomic coverage and resolution. The most numerous unique taxa were identified using 18S-
43 dinoflagellate amplicon (N=167) and the best resolution to species level occurred using LSU
44 (N=60). This work is the first report of HAB species identification in Canada using a
45 combination of morphological, metabarcoding, and qPCR approaches. These results emphasize
46 the benefit of applying molecular techniques to detect HAB taxa and highlight the current
47 necessity of using multiple markers for accurate detection of the diverse groups that cause
48 HABs.

49

50

51 **Keywords**: Harmful algal blooms, Next-Generation Sequencing, Metabarcoding, *Alexandrium*,
52 *Heterosigma*, *Pseudochattonella*

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

53 # INTRODUCTION

54 Phytoplankton form the base of the marine food web and are required to support healthy aquatic
55 ecosystems. In some circumstances, however, high-biomass events and/or proliferation of certain
56 algal species can cause harm to aquatic animals through a variety of means including disruption
57 of the food web, shellfish poisoning, the development of low oxygen 'dead zones' after bloom
58 degradation, and fish kills through toxins, gill damage, or hypoxia (Rensel and Whyte, 2004).
59 Collectively, these events are termed harmful algal blooms (HABs). Importantly, there is a
60 general scientific consensus that public health, fisheries, and ecosystem impacts from HABs
61 have all increased over the past few decades (*e.g.* Andersen, 2012; Hallegraeff, 2004).

62 The coastal waters of British Columbia (BC), Canada, in the northeast Pacific Ocean,
63 have one of the longest documented histories of severe HABs going back to the first reported
64 case in 1793 (Vancouver, 1798). A government program for monitoring the presence of toxins in
65 shellfish was established in the early 1940s (Taylor and Harrison, 2002) and since then paralytic
66 shellfish poisoning (PSP) closures have occurred every year. The BC salmon aquaculture
67 industry initiated and has been supporting the Harmful Algae Monitoring Program (HAMP)
68 since the 1990s due to the devastating effects of harmful algae on farmed fish (Horner et al.,
69 1997; Rensel and Whyte, 2004). During 2009–2012, direct losses to the BC salmon aquaculture
70 industry from HABs were ~13 M USD (Haigh and Esenkulova, 2014). HABs are currently one
71 of the most significant risks for the BC aquaculture industry, regularly causing severe economic
72 losses through finfish/shellfish mortalities and shellfish harvest closures due to toxin
73 accumulation (Whyte et al., 1997). Therefore, there is an ongoing and pressing need for
74 monitoring and research on HABs phenomena in coastal BC.

75 Monitoring HABs typically depends on the effective identification and enumeration of
76 species of concern in water samples. Algal cell identification has long been accomplished based
77 on morphology revealed through visual microscopic examination. Although traditional light
78 microscopy currently remains the standard, it has limitations when it comes to certain species
79 and strains that cannot be easily visualized or cannot be discriminated between harmful and
80 benign variants based on morphological characteristics alone (Hallegraeff, 2004). Moreover,
81 microscopic identification is highly dependent on the level of expertise and experience of the
82 individual analyzing the samples and with fewer morphological taxonomists being trained, it is
83 increasingly difficult to keep up with the demand. In recent years, studying HABs with

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

84 molecular techniques, either in tandem with morphological methods or independently, has
85 become increasingly popular. Quantitative real-time polymerase chain reaction (qPCR) is a
86 powerful method for detecting and quantifying DNA over a broad dynamic range (Livak and
87 Schmittgen, 2001). This method has been used for species-specific harmful algal detection and
88 enumeration (*e.g.* Antonella and Luca, 2013; Eckford-Soper and Daugbjerg, 2015; Scholin et al.,
89 2011). High-throughput platforms, such as the Fluidigm BioMark, allow for the use of multiple
90 species-specific probes for simultaneous detection and enumeration of multiple taxa (Medlin and
91 Orozco, 2017) with the potential for both time and cost savings, as well as the additional ability
92 to identify cryptic species, compared to traditional light microscopy. Next-generation sequencing
93 (NGS) methods, as applied through metabarcoding, allow for millions of sequencing reactions to
94 be performed in parallel, resulting in the ability to generate massive amounts of sequencing data
95 (Goodwin et al., 2016; Valentini et al., 2016). Sequence-based taxonomic approaches can allow
96 for the identification of multiple species of interest, including nano- and picoplankton, rare and
97 fragile taxa, and cryptic species, in a reproducible and cost effective manner (*e.g.* Eiler et al.,
98 2013).

99 The objective of the present study was to obtain microalgal taxa that are known or
100 suspected to be harmful to cultured fish and shellfish in BC and to identify these taxa through
101 light microscopy as well as genetic methods. Fish-killing algae targeted included *Chaetoceros*
102 *concavicorne*, *C. convolutus*, *Chattonella* sp., *Chrysochromulina* spp., *Cochlodinium fulvescens*,
103 *Dictyocha* spp., *Heterosigma akashiwo*, *Karenia mikimotoi*, and *Pseudochattonella verruculosa.*
104 Shellfish-poisoning algae included *Alexandrium* spp., *Dinophysis* spp., *Protoceratium*
105 *reticulatum*, and *Pseudo-nitzschia* spp. During this study, cultures and water samples for
106 identification of these species were primarily acquired from HABs occurring in coastal BC. Here
107 we describe, compare, and cross validate coastal BC harmful algae identification based on
108 morphology, qPCR, and metabarcoding.

109

110 # METHODS

111 ***Ethics Statement***

112 No permits were required for collection of water samples in Canadian coastal waters.

113

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

114 ***Sample Collection***

115 All samples used in this study were collected in coastal BC waters (Table 1), except for two that
116 were obtained from the Provasoli-Guillard National Center for Marine Algae and Microbiota
117 (Samples s01, s02) that were used to test primer specificity and to optimize qPCR reactions. All
118 samples (N=54), with their respective sampling locations, are summarized in Table 1 and Fig. 1.

119 Field samples of algae were obtained by collecting seawater with a 1 L sampling bottle
120 (Venrick, 1978). Samples to be used for culturing and genetic analyses were kept cool and dark
121 until processed. Samples for microscopic identification were immediately preserved with
122 Lugol's iodine (Andersen and Throndsen, 2004) and then shipped to the laboratory for
123 taxonomic analysis. Identification and enumeration of phytoplankton based on morphology was
124 done with a compound microscope using a Sedgewick-Rafter slide (Guillard, 1978).
125 Identification based on morphology was done to the lowest possible taxonomic level (Hasle,
126 1978), by an experienced phytoplankton taxonomist.

127

128 ***Cultures***

129 Cultures for this study were isolated by serial dilution from the unpreserved field water samples
130 and from sediments collected from sites with HABs observed in the past (see Table 1 for sample
131 origins), as some harmful algae produce cysts that play an important role in initiating subsequent
132 blooms (Anderson et al., 2003). Several sediment samples were taken from three areas around
133 Vancouver Island (*i.e.* Baynes Sound, Okeover Inlet, Quatsino Sound; Fig. 1) in late winter and
134 spring 2013. Target species for isolation and germination were *Alexandrium* spp. due to the
135 numerous PSP closures recorded in Baynes Sound and Okeover Inlet in summer 2012 (DFO,
136 2012), as well as *H. akashiwo* and suspected *Chattonella* spp. that caused fish mortalities in late
137 summer 2012 in Quatsino Sound (Haigh, personal observation). Sampling was done using van
138 Veen and Ekman grab samplers. Sediment samples (3 – 5 L) were taken from depths of 5 – 45
139 m. Sites were chosen based on the assumption that cysts accumulate in the same areas as silt and
140 clay (Wall and Dale, 1967). The upper 3 cm of sediments in the grab samples were collected,
141 mixed well, and immediately placed in containers and kept cool and dark.

142 Culture establishment from the collected samples occurred as described in detail
143 previously (Esenkulova et al., 2015). In brief, sediments were sieved to produce a 20 – 120-μm
144 fraction and incubated in enriched natural seawater medium (Harrison et al., 1980) at 18 °C

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

145 under continuous illumination (20 µmol photons m^{-2} s^{-1} provided by full-spectrum fluorescent
146 lamps). Serial dilutions were performed in a laminar flow hood using a micropipette and sterile
147 24-well plates. Several cultures of *A. tamarense*, *H. akashiwo*, and *Chattonella* spp.? (question
148 mark indicates that taxon is suspected, but not positively identified by light microscopy) were
149 successfully established. Isolation from water samples produced cultures of *Alexandrium* sp.,
150 *Chaetoceros eibenii* (a non-harmful species), *Chattonella* sp.?, *H. akashiwo*, *Protoceratium*
151 *reticulatum*, *Pseudochattonella* cf. *verruculosa*, and *Pseudo-nitzschia* spp. Some cultures were
152 pure (appeared to contain a single species) and some were mixed (contained multiple species).
153
154 ***DNA Extraction and Purification***
155 Water samples for genetic analyses were either filtered or pelleted. For filtered water samples,
156 60–100 mL of water was filtered through 25-mm or 47-mm GF/F Whatman filters and then
157 frozen at -20 °C or stored in 75–95% ethanol until DNA extraction. A section of ¼ of the 47-mm
158 filters or ½ of the 25-mm filters were used for the extraction. Water samples concentrated by
159 centrifugation were subsampled and 1.5 – 40 mL centrifuged at 8,000 x g, the supernatant
160 removed, and the resultant pellet stored in 95% ethanol until DNA extraction. The variation in
161 sampling volumes and sample methods occurred due to the large variation of suggested methods
162 in the available literature. Cultured samples for molecular analyses were taken from ~200-mL
163 cultures in a laminar flow hood, whereby samples of well-mixed cultures were divided in 2 x 20-
164 mL subsamples and centrifuged for 15 min at 5,000 x g for 10 min. The supernatant was
165 decanted, samples were re-suspended in 1 mL of phosphate buffered saline, transferred to 2-mL
166 microfuge tubes, and then centrifuged again at 4,000 g for 10 min. Finally, the supernatant was
167 again removed and cell pellets frozen at -20 °C until extraction.
168 DNA was extracted and purified from all samples using the DNeasy Blood and Tissue kit
169 (Qiagen, Toronto, Canada) as per manufacturer's instructions, with the addition of a
170 homogenization step using a TissueLyser 2 homogenizer (Qiagen) with a 4-mm steel bead. DNA
171 concentration was measured by spectrophotometry (NanoDrop, ND-1000).
172
173 ***qPCR Assays***
174 A literature search was undertaken to identify published TaqMan-based qPCR assays for harmful
175 algal taxa of interest in the northeast Pacific Ocean. TaqMan assays were found for five targeted

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

176 taxa: six assays for *Alexandrium* spp., two for *Chattonella* spp., three for *H. akashiwo,* two for *K.*
177 *mikimotoi*, and two for *P. verruculosa*. Additional published TaqMan assays for known harmful
178 algae from other parts of the world were also included, bringing the total number of target algal
179 taxa to 28 and the number of assays to 39 (Table 2).

180 The qPCR reactions using all assays (Table 2) were conducted on a Fluidigm BioMark™
181 platform (Fluidigm Corporation, San Francisco, CA, USA) using a 96x96 dynamic array to run
182 9,216 reactions simultaneously (96 samples with 96 assays) as described in detail in Miller et al.
183 (2016). To reduce the effect of PCR inhibitors that can be problematic in algae, samples were
184 tested at concentrations of 10 and 2.5 ng uL^{-1} or, if the DNA concentration was below that, at the
185 highest available concentration. Each assay was run in duplicate.

186 Briefly, a pre-amplification (STA) step of 14 cycles using dilute (50 nM in a 5-uL
187 reaction) primer pairs of each assay with TaqMan Preamp MasterMix (Applied Biosystems,
188 Foster City, CA, USA) was performed according to the BioMark protocol (Applied Biosystems).
189 Unincorporated primers were removed with ExoSAP-IT (Affymetrix, Santa Clara, CA, USA)
190 and samples were then diluted 1:5 in DNA Suspension Buffer (Teknova, Hollister, CA, USA). A
191 5-µL sample mix was prepared for each pre-amplified sample with TaqMan Universal Master
192 Mix (Life Technologies Corporation, Carlsbad, CA, USA) and GE Sample Loading Reagent
193 (Fluidigm Corporation) and a 5-µL aliquot of assay mix was prepared containing 10-µM primers
194 and 3-µM probes for each separate TaqMan assay and each was loaded onto a pre-primed
195 dynamic array. An IFC controller HX pressurized and mixed the assays and samples from their
196 individual inlets on the dynamic array and the PCR was run on the BioMark with the following
197 conditions: 50 °C for 2 min, 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C
198 for 1 min. Output data was analyzed and the cycle threshold (Ct) per sample determined using
199 Fluidigm Real Time PCR Analysis software (Fluidigm Corporation). Cycle threshold values for
200 qPCR replicates were averaged in the final results.

201

202 ***Metabarcoding and Sequencing***

203 Four primer pairs were selected for use in NGS metabarcoding from published studies with the
204 goal of amplifying a broad range of algal taxa (see Table 2). Illumina adaptors were incorporated
205 onto the 5' end of each primer for the attachment of Nextera XT Illumina indices (Illumina, Inc.,
206 San Diego, CA, USA) in the second round of PCR. Input sample DNA was normalized to 5 ng

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

207 μL^{-1} when starting concentrations allowed. The primer sets were used as the initial primers in a
208 two-step PCR process to construct and sequence four Nextera XT libraries according to the
209 Illumina 16S Metagenomic Sequencing Library Preparation protocol (15044223, Rev. B).
210 Library quantification, normalization, pooling, and sequencing on a MiSeq with a 600-cycle flow
211 cell (MiSeq Sequencing Kit v3, 600 bp, Illumina, Inc.) were performed according to the
212 manufacturer's protocols, with the only modification being that the final, pooled library was run
213 at a concentration of 16 pM with 10% PhiX control.
214

215 ***Bioinformatics***

216 All sequence data was de-multiplexed using input sample barcodes during file export from the
217 sequencer (Illumina, Inc.), resulting in a pair of fastq files for each individual sample. Quality of
218 raw sequence fastq files was evaluated in FastQC (Andrews and Babraham Bioinformatics,
219 2010) with results aggregated using MultiQC (Ewels et al., 2016). In general, the OBITools
220 package (Boyer et al., 2016) was used for the analysis of the different amplicons, but each
221 amplicon required specific inputs due to different features of the data. All bioinformatics steps
222 are outlined in detail on GitHub (see *Data Availability*), and explained here per marker type (*i.e.*
223 LSU, 18S, 16S). For LSU and 18S, due to the larger size of amplicons and the lack of read
224 merging for these paired-end datasets, only the forward reads were used. For 16S, paired-end
225 data were used as most reads overlapped and therefore were able to be merged.

226 For LSU (single-end data), primers were removed using cutadapt (Martin, 2011) and then
227 reads were formatted for the OBITools pipeline using *ngsfilter* (OBITools) on each sample
228 without any input barcodes. Each formatted input sample file was then annotated with a sample
229 identifier in read headers with *obiannotate*. Sample labeled fastq files were then combined to a
230 single file, with reads cut to a uniform size (*i.e.* 230 bp) to reduce singleton records due to slight
231 differences in read length for single-end data. Subsequently, this file was moved into the
232 standard OBITools pipeline (see below).

233 For 18S (single-end data), there were two different amplicons, one for dinoflagellates and
234 one for diatoms (*i.e.* 18S-dinoflagellate, 18S-diatom). These two amplicon types were de-
235 multiplexed using *ngsfilter*, which output a single fastq file per sample, but from both amplicon
236 types. A custom script separated the two types of amplicons (see *Data Availability*). Obiannotate
237 was used to annotate each sample-amplicon type with sample and amplicon identifiers. All data

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

238 was then merged into a single file, cut to a uniform size using cutadapt (as above), and then
239 moved into the standard OBITools pipeline.

240 For 16S (paired-end data), primers were removed using cutadapt, overlapping reads were
241 merged using *illuminapairedend* retaining merged data with an overlap score ≥ 40 (OBITools),
242 samples were annotated with a sample identifier, and then combined to a single file and moved
243 into the standard OBITools pipeline.

244 Once all data were input in a uniform format into the main pipeline as described above,
245 the OBITools package was used to retain a single representative accession per unique amplicon
246 sequence, keeping record of the number of reads per sample for the accession in the accession
247 header using *obiuniq*. Subsequently, the data were de-noised using a size filter and a low count
248 threshold (*obigrep*) and by removing 'internal' sequences (*i.e.* probable PCR/sequencing errors;
249 *obiclean*), as per standard OBITools approaches (Boyer et al., 2016). The counts per
250 representative unique amplicon were exported using *obitab*.

251 The unique amplicon file for each amplicon type was annotated using BLAST (Altschul
252 et al., 1997) run in parallel (Tange, 2011), receiving ten alignments per BLAST query. For each
253 unique amplicon accession, taxa were assigned using MEGAN (Huson et al., 2016) using the
254 Lowest Common Ancestor (LCA) algorithm with the following parameters: min score = 100;
255 max expected = 10^{-9}; min % ID = 97; top % 10; min support % (off); and min support = 1. Any
256 amplicon that received a BLAST result, but was not assigned using MEGAN due to parameters
257 set within MEGAN was put in the 'Not assigned' category. Any amplicon without a BLAST
258 result at all was put in the 'Unknown' category. Read counts were connected to annotations from
259 the MEGAN output using custom R ("R Core Team.," 2018) scripts (see *Data Availability*). A
260 threshold of at least 10 reads per sample-taxon combination was applied to reduce the potential
261 for false positive detection and any sample-taxon combination with fewer than this was
262 transformed to 0. Taxa proportions were calculated by dividing the count per taxon for a sample
263 by the total number of reads for the sample. Taxon ranks and the classification of identified taxa
264 was enabled using *taxize* (Chamberlain and Szöcs, 2013; Scott et al., 2019) in R, and a custom
265 database was created to contain assembly taxonomy, read count, and sample information. Bar
266 plots were constructed using ggplot2 in R (Wickham, 2016) and pie charts constructed using
267 Krona (Ondov et al., 2011).

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

268 # RESULTS

269 ***Morphological Identification***

270 In total, 54 samples containing either cultured (N=30) or field (N=24) collections of ten
271 suspected harmful algal taxa identified by morphology were obtained during this study,
272 including *Alexandrium* spp., *Chattonella* sp., *Chrysochromulina* spp., *Cochlodinium fulvescens*,
273 *Dictyocha* spp., *Heterosigma akashiwo*, *Protoceratium reticulatum*, *Pseudochattonella*
274 *verruculosa*, *Pseudo-nitzschia* spp., and *Pseudopedinella* sp. (note: *Pseudopedinella* was not
275 initially targeted, but was acquired opportunistically at a suspected HAB event). Although other
276 species were initially targeted, they were unable to be acquired for complete analysis, due to
277 either low cell concentrations in field samples (*i.e. C. concavicornis*, *C. convolutus,* and
278 *Dinophysis* spp.) or due to a general absence in field samples (*i.e. K. mikimotoi*). All samples
279 obtained for this study are listed in Table 1.

280

281 ***Extraction Efficiency for Molecular Methods***

282 Sample collection volumes, types, and preservation methods varied throughout the collection
283 time period as sampling and extraction protocols were still being optimized. As such, DNA
284 quantity in extractions was highly variable due to the variety in the sample state, density, and
285 preservation methods applied. Cultured samples gave the highest yields. Filtered field samples
286 generally resulted in lower, but adequate yields. Pelleted field samples resulted in low and often
287 unusable yields. After extractions, 46 of the 54 samples had sufficient high-quality DNA for
288 molecular analyses. Samples with insufficient DNA (*i.e.* s025, s026, s027, s029, s039, s051,
289 s052, s054) were not run on qPCR, but two of these samples (*i.e.* s029 and s054) were run
290 through NGS and returned OTUs (<25 and <1,350, respectively). Five samples (*i.e.* s034, s045,
291 s046, s050, s053) had low DNA concentration, but returned positive results with qPCR and
292 NGS.

293

294 ***Taxa Identified by qPCR***

295 A total of 46 samples with sufficient DNA concentrations were run using qPCR. However, only
296 eight of 39 applied TaqMan assays provided amplification with one or more of the samples
297 (Table 3), and the rest (N=31 assays) did not return positive results in any samples. These eight
298 assays amplified within 44 of the samples. Targeted species that were detected via qPCR

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

299 included *Alexandrium* spp./*A. tamarense*, *H. akashiwo*, and *P. verruculosa*. The targeted
300 *Chattonella* spp. had four different qPCR assays, but none resulted in detections. Other species
301 that are potentially harmful were identified by qPCR, but were not specifically identified by
302 microscopy. For example, cyanobacteria species were detected in nine samples, *G. instriatum* in
303 three and *Karlodinium micrum/veneficum* in eleven samples. Nine targeted taxa were not
304 quantifiable by qPCR due to a lack of published TaqMan assays: *C. concavicornis*, *C.*
305 *convolutus, Chrysochromulina* spp., *Cochlodinium fulvescens*, *Dictyocha* spp., *Dinophysis* spp.,
306 *P. reticulatum*, and *Pseudo-nitzschia* spp. Due to the fact that this study ran the assays under
307 different conditions than they were designed for, and due to algae strain variability, the
308 possibility for false negatives or positives cannot be excluded. To confirm if the sample
309 contained a target species, and therefore to determine the effectiveness of the tested assays and
310 correspondence to microscopy, NGS was conducted. The relative effectiveness of the qPCR
311 results in comparison to microscopy and metabarcoding is provided in the *Discussion*.
312

313 ***Metabarcoding Overview***

314 In total, over 350 individual taxa were detected by metabarcoding and NGS. NGS returned
315 results for 48 of 48 sequenced samples. Detailed results of the analysis including all taxa and
316 associated read counts are provided in Supplemental Tables S1-4.

317 The most numerous unique taxa occurred through using the 18S-dinoflagellate amplicon
318 (Table 4), which returned 167 individual taxa, including 49 to the species level. For comparison,
319 the 16S amplicon identified 69 individual taxa (21 to the species level), and 18S-diatom 78 taxa
320 (18 to the species level). The LSU data identified 136 individual taxa and was the most effective
321 at resolving data to the species level (N=60 species), which is noteworthy given the importance
322 of species-level identification when identifying harmful algae. Only one taxon, *H. akashiwo*, was
323 identified by all four amplicons. Many similar taxa (N=20) were detected by 18S-diatom, 18S-
324 dinoflagellate, and LSU amplicons, but only three were identified to the species level.

325 The numbers of reads per taxonomic category, with an emphasis on microalgal groups,
326 are listed in Table 5. Total microalgae reads detected by 18S-diatom, 18S-dinoflagellate, and
327 LSU runs were more than 85% of all reads (not including unknowns) for each amplicon, whereas
328 microalgae reads from 16S run comprised less than 50%. The majority of microalgal reads for all
329 amplicons belonged to raphidophytes. The 18S-dinoflagellate amplicon detected the most

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

330 microalgal groups and the 16S amplicon detected the fewest microalgal groups. Silicoflagellates
331 (Dictyochophyceae), an important group that includes several harmful and potentially harmful
332 species, were detected only by 18S-diatom and 18S-dinoflagellate amplicons.

333 All harmful and potentially harmful algal taxa reads (*i.e.* at the species, genus, and family
334 level) detected by all amplicons are provided in Table 6. For a more convenient comparison
335 between results employing different amplicons, grouping of harmful and potentially harmful
336 algae species to genus level was done for *Alexandrium* spp., *Chattonella* spp., *Dinophysis* spp.,
337 *Karlodinium* spp., *Phalacroma* spp., *Prymnesium* spp., *Pseudo-nitzschia* spp., and
338 *Pseudochattonella* spp.

339 Overall, most of the reads of harmful and potentially harmful algae were detected by
340 LSU and 18S-dinoflagellate amplicons (Table 6). At the species and genus levels, the LSU
341 amplicon provided the most reads for dinoflagellate species within the *Cochlodinium*,
342 *Dinophysis*, *Karlodinium*, and *Phalacroma* genera, as well as haptophytes in the
343 *Chrysochromulina* and *Prymnesium* genera. The 18S-dinoflagellate amplicon detected the most
344 reads for species within the *Alexandrium*, *Pseudochattonella*, and *P. reticulatum* genera, whereas
345 the 18S-diatom amplicon detected the most reads for *Pseudo-nitzschia* spp. as well as the
346 raphidophytes *Chattonella* sp. and *H. akashiwo*. There were no instances when 16S detected the
347 most reads for any of the listed taxa.

348 The total reads and percentages per sample detected by different amplicons are shown in
349 Figure 2. Specifically in this plot, when a taxon of harmful or potentially harmful algae had less
350 than 100 reads, it was grouped within an appropriate, larger algae category, *e.g. Karlodinium*
351 spp. with 77 reads were included in the dinoflagellates counts; *Prymnesium* spp. with 20 reads
352 and Prymnesiaceae with 11 reads were included in haptophytes (Figure 2). Overall, based on the
353 results for all four amplicons, culture samples had a higher number of reads, by an order of
354 magnitude, and lower algal species diversity than field samples (Figure 2, Supplement 5).

355 A comparison of microscopy and NGS taxonomic identification in each of the 48 samples
356 revealed that the majority of the taxa detected by microscopy also were identified in the NGS
357 results (Supplement 6). Species and genera that were positively identified by both microscopy
358 and NGS included: *Alexandrium*, *Chaetoceros, Cochlodinium, H. akashiwo, P. reticulatum,*
359 *Pseudochattonella*, and *Pseudo-nitzschia*. There were a few notable exceptions, with the two
360 most consistently observed mismatches being (i) all suspected *Chattonella* spp. (10 field and

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

361 cultures) were identified by NGS as *H. akashiwo*; and (ii) *Dictyocha* spp., *P. verruculosa*,
362 *Pseudochattonella* sp., and *Pseudopedinella* spp. (*i.e.* all three silicoflagellate genera) were not
363 detected in 11 out of 15 samples by any of the amplicons. Although these samples had unknown
364 and unassigned reads in NGS results (Supplement 6), the sequences for all of the species not
365 identified by NGS but identified with microscopy were in fact present in the Gen Bank database,
366 and therefore the reason for the lack of these taxa, as well as the identity of these unknown reads
367 remains unknown.

368

369 # DISCUSSION

370 ***Harmful Algae: Comparing Microscopy and Molecular Approaches***

371 The term 'algae' comprises a diverse, polyphyletic group that encompasses organisms from
372 widely different taxonomic domains and includes eukaryotes and prokaryotes. Major algal
373 groups include diatoms, dinoflagellates, raphidophytes, silicoflagellates, haptophytes, and
374 cyanobacteria. Identification methods applied to each of these groups will be outlined in detail
375 below.

376

377 ***Diatoms***

378 Diatoms are the major group of microalgae with approximately 250 modern genera and
379 thousands of species (Hasle and Syvertesen, 1996). There are only a few diatom taxa that are
380 harmful due to toxin production (*i.e.* several species from the *Pseudo-nitzschia* genus) or cell
381 structures that cause mechanical damage to gills (*e.g. C. concavicornis* and *C. convolutus*). In
382 our study, only one sample contained a potentially harmful diatom – the culture of *Pseudo-*
383 *nitzschia* sp. Positive identification of *Pseudo-nitzschia* to the species level with microscopy can
384 be ensured only with scanning electron microscopy (SEM), which is a very costly and time-
385 consuming technique not generally implemented in routine monitoring. In our study, both the
386 18S-diatom and LSU amplicons showed promising results by confirming *Pseudo-nitzschia*
387 presence (>10,000 reads, >98% of the total reads in the cultured sample). However, it also did
388 not discriminate to the species level. The same amplicons detected very low presence (<1%) of
389 *Pseudo-nitzschia* spp. in five other samples where cell presence was not detected using
390 microscopy. Most of these samples were presumed monocultures of other species that were

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

391 processed on the same day (*i.e.* culture subsampled, DNA extracted and purified), and NGS
392 results may indicate possible low-level cross contamination from high concentration, pure
393 cultures that occurred in the lab in the same flow hood.
394
395 ***Dinoflagellates***
396 Dinoflagellates are second to diatoms in terms of importance to marine primary production, with
397 about 2000 described extant species (Taylor et al., 2008). This group is particularly important for
398 HABs, as about 75–80% of known toxic phytoplankton species belong to this group (Cembella,
399 2003). Microscopy indicated that there were nine samples containing *Alexandrium* spp.,
400 *Cochlodinium* sp., or *P. reticulatum*. However, based on molecular techniques, additional
401 potentially harmful dinoflagellate species were detected (*Dinophysis*, *Karlodinium*, and
402 *Phalacroma*) and considerably higher numbers of samples containing them were identified. For
403 dinoflagellates, as expected, the most conclusive results were obtained using the 18S-
404 dinoflagellate and LSU amplicons.
405
406 ***Dinoflagellates: Alexandrium***
407 Many species of *Alexandrium* produce toxins that cause paralytic shellfish poisoning (PSP). In
408 BC, since the establishment of the monitoring program for toxins in shellfish in the 1940s
409 (Taylor and Harrison, 2002), PSP closures have occurred every year, and this negatively affects
410 shellfish aquaculture and recreational harvesting. Six cultured samples of *Alexandrium* were
411 included in this study. One was a purchased culture of *Alexandrium* cf. *catenella* and five were
412 locally established cultures microscopically identified as *Alexandrium* sp. and *A. tamarense;* four
413 of these cultures were monocultures and one was a mixed culture that also contained green algae
414 and flagellates. The most effective metabarcoding results for *Alexandrium* detection were
415 obtained using the 18S-dinoflagellate amplicon, closely followed by LSU. Both amplicons
416 provided comparable results with a very high percentage of total *Alexandrium* reads per sample
417 for all five of the culture samples (>90% of reads on average) and a low percentage in the mixed
418 culture (<4%). All six of these samples had highly positive qPCR results using an assay targeting
419 the toxic North American strain of *A. tamarense* (Toebe et al., 2013).
420 Detections of *Alexandrium* in samples where it was not identified based on microscopy
421 were less conclusive. There were 19 samples where *Alexandrium* presence was suggested by

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

422 18S-dinoflagellate and LSU amplicons but not by microscopy. In all cases these were very low
423 counts of *Alexandrium* (*i.e.* very low percent of reads per sample, ranging from 0.02% to 4.72%).
424 However, six of these 19 samples also had positive qPCR results with the *A. tamarense* assay,
425 suggesting that the positives were not completely attributable to false positives from
426 metabarcoding methods. In addition, there were 10 samples with *A. tamarense* qPCR detection
427 that were not confirmed by microscopy or metabarcoding. Future work is needed to clarify
428 exactly which results are correct, the microscopy or the molecular (qPCR and metabarcoding) for
429 this taxon, and how to best score these detections (*e.g.* weak positives). With the expected higher
430 sensitivity of molecular assays as compared to microscopy, it is not implausible that low copy
431 numbers of a HAB species may exist in many mixed cultures, but whether these were introduced
432 by cross-contamination of cultures, as well as the biological relevance of such levels also would
433 need further study.

434 While *Alexandrium* presence detected by 18S-dinoflagellate and LSU amplicons was
435 similar, the biggest difference in these amplicons was in the taxonomic level of identification. In
436 the 18S-dinoflagellate amplicon, more than 99.50% of *Alexandrium* OTUs were identified to
437 genus level only, while the very small remainder was assigned to *A. tamarense* (0.46%) and *A.*
438 *andersonii* (<0.01%). In contrast, the LSU results indicated that 97.31% of total *Alexandrium*
439 reads belonged to the species *A. fundyense*, 0.05% to *A. tamarense*, with the remainder 2.64%
440 assigned to the genus level. Based on morphological characteristics, three species (*A. catenella*,
441 *A. tamarense*, and *A. fundyense*) comprise the *A. tamarense* species complex (Balech, 1985).
442 These morphospecies are quite similar in appearance and are distinguished mainly based on
443 formation of chains of cells and features of the cell theca: presence and shape of a ventral pore
444 between plates 1’ and 4’, and differences in the shapes of plates. However, recent rDNA
445 sequencing studies (John et al., 2014; Lilly et al., 2007) found that these three morphospecies did
446 not satisfy requirements for biological species definition. They suggested that these groups be re-
447 evaluated and a new species definition was proposed. In our study, clonal cultures that were
448 morphologically identified as *A. catenella*, *A. tamarense*, and *Alexandrium* spp. were all assigned
449 to *A. tamarense* by TaqMan and to *A. tamarense* or *A. fundyense* species by metabarcoding,
450 depending on the amplicon. This ambiguity further supports the proposition that the taxonomy of
451 the *A. tamarense* species complex should be re-evaluated, with consensus classifications by
452 molecular methods needing to be updated in sequence databases (*e.g.* GenBank).

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Dinoflagellates: Cochlodinium

Two out of approximately 40 known *Cochlodinium* species (*i.e. C. polykrikoides* and *C. fulvescens*) form HABs and can cause fish kills (Kudela and Gobler, 2012). In BC, blooms of *C. fulvescens* were implicated in farmed salmon kills that caused ~1.5 M USD in economic losses (Whyte et al., 2001). Here, *Cochlodinium* was observed by microscopy in two field samples. Its presence (24 and 992 reads) was detected at the genus level in both of these field samples, but only by the LSU amplicon. This identification was very conclusive and OTUs of *Cochlodinium* at the species or genus level were not observed in any of the other 44 samples.

Dinoflagellates: Protoceratium reticulatum

Protoceratium reticulatum produces yessotoxins, which can be bioaccumulated by shellfish and have been associated with diarrhetic shellfish poisoning (DSP) (Satake et al., 1998). Metabarcoding was very effective for identification of this species. The sample of *P. reticulatum* from an established culture with identification based on morphology had a positive identification to the species level when using both the 18S-dinoflagellate and LSU amplicons (143,019, 42,069 reads). The 18S-dinoflagellate amplicon also detected very low presence (<2% of the total reads per sample) of *P. reticulatum* in another three culture samples and one field sample (s34). Although the LSU amplicon had comparable readings and percentages to the 18S-dinoflagellate amplicon, it had lower detections in the unexpected samples (*i.e.* two out of four samples were positive by LSU).

Dinoflagellates: Dinophysis*, *Gymnodinium*, *Karlodinium*, and *Phalacroma

Algae from the *Dinophysis*, *Gymnodinium*, *Karlodinium*, or *Phalacroma* genera were not identified in the samples using microscopy, but were detected using molecular methods. Certain *Dinophysis* and *Phalacroma* species produce toxins that cause DSP, and thus are important to shellfish aquaculture. In BC, the first reported DSP outbreak was associated with elevated numbers of *Dinophysis* spp. (Esenkulova and Haigh, 2012). In the present study, a low number of reads (<200 in total) of *D. parvula*, *P. rapa* (previously known as *D. rapa*), as well as *Dinophysis* spp. were detected in one field sample using the LSU amplicon. The *Karlodinium* genus includes several toxin-producing species; for example, *K. veneficum* blooms have been

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

484 associated with aquatic faunal mortalities for decades (Place et al., 2012). A small number of
485 reads to the genus and species level (<100 in total) were detected in one field sample by LSU.
486 Similar results for both *Dinophysis* and *Karlodinium* were obtained using the 18S-dinoflagellate
487 amplicon, but were only annotated to the genus level. However, metabarcoding results did not
488 confirm *Karlodinium* presence in ten other samples with positive qPCR results (range: 14 to 29
489 Ct).

490 Some *Gymnodinium* species produce toxins (Moestrup et al., 2008) and in the present
491 study there were very weak detections of *G. instriatum* in three samples by qPCR (Ct >30), but
492 none were found positive by metabarcoding at the species, genus, or order level.

493

494 ***Raphidophytes***

495 Raphidophytes are a group of algae with very few species. However, these species include
496 several taxa that pose some of the most serious threats to finfish aquaculture around the world,
497 such as *H. akashiwo* and those in the *Chattonella* genus (Hallegraeff, 2004). In our work, the
498 best results for raphidophytes were obtained using 18S-diatom and LSU amplicons.

499

500 ***Raphidophytes: Heterosigma akashiwo***

501 *Heterosigma akashiwo* is a major fish killer in BC, causing economic losses to the BC salmon
502 aquaculture industry of about ~3.5 M USD per year (Haigh and Esenkulova, 2014). All 24
503 samples (18 cultures and 6 field) where *H. akashiwo* presence was observed with microscopy
504 returned positive detections with qPCR and with three of the amplicons. The amplicon with
505 weaker detection ability was the 18S-dinoflagellate, which identified *H. akashiwo* in only 19 out
506 of 24 samples. The highest metabarcoding reads were detected by the 18S-diatom amplicon. In
507 the rest of the samples (N=24), *H. akashiwo* presence was not noted by microscopy observations,
508 but *H. akashiwo* reads were found in more than 20 of these samples by the 18S-diatom, 16S, and
509 LSU amplicons, in six using the 18S-dinoflagellate amplicon, and in 11 of these samples by
510 qPCR. For most of the samples there was a general agreement between *H. akashiwo* read
511 numbers by the different techniques (*e.g.* all 18 samples where *H. akashiwo* was not detected by
512 18S-dinoflagellate had <100 reads per sample by all other techniques; most of the 15 samples
513 with the highest *H. akashiwo* reads detected by 18S-diatom had the highest levels with other
514 techniques). One exception to this trend was in the purchased culture of *Chattonella subsalsa*,

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

515 where a high load of *H. akashiwo* was identified solely by the 16S amplicon. *C. subsalsa*
516 identification in this case was confirmed by microscopy and other amplicons, so this clearly
517 indicates a 16S database issue (*i.e.* mis-representation of *C. subsalsa* species as *H. akashiwo*). In
518 35 samples where *H. akashiwo* presence was suggested based on qPCR results, there was a
519 general agreement of higher loads (*e.g.* Ct<20) with high OTU read counts (>4,000) in the 18S-
520 diatom amplicon, but not in other amplicons. However, some samples with lower qPCR loads
521 (CT=20–29.5) also had very high OTU read counts, with nine samples containing over 10,000
522 reads. Therefore the quantitation between the read counts and the qPCR was not always
523 congruent. It is possible that this could be due to mismatches in specificity for the applied assay,
524 but further work would be needed to better understand this difference.
525
526 ***Raphidophytes: Chattonella***
527 Globally, many species of *Chattonella* have been associated with fish kills (Moestrup et al.,
528 2008) and blooms of *Chattonella* sp. have caused farmed fish mortalities in BC (Haigh and
529 Esenkulova, 2014). Here, all ten local samples with positive microscopy identification of
530 *Chattonella* spp. did not return results for this genus by sequencing. Almost all (>99.99%) of the
531 Raphidophytes reads in these ten samples (*i.e.* one culture, five mixed cultures, and four field
532 samples) where *Chattonella* presence was suspected based on microscopy, were assigned to *H.*
533 *akashiwo* by all four amplicons. It is possible that large *Heterosigma* cells were misidentified as
534 *Chattonella* during microscopy. The purchased culture of *Chattonella subsalsa* was positively
535 confirmed as *Chattonella* (>99% of total reads) by 18S-diatom and LSU amplicons with a very
536 small portion (0.04%) identified to the species level (*C. subsalsa*) using the LSU amplicon,
537 suggesting that the amplicons could identify *Chattonella* if it was present. Therefore, for this
538 species the NGS approach appeared to be superior to light microscopy identification.
539 A very low number of reads (<30) of *Chattonella* spp. were also detected in four culture
540 samples of *Alexandrium* spp. and *H. akashiwo* by the 18S-diatom amplicon. This small number
541 is most likely also an artefact of the subsampling process (*see above*). Only a small portion of the
542 purchased *C. subsalsa* culture was identified to the species level, and the rest of the non-
543 purchased samples was identified to genus only, which suggests the difference in resolution may
544 be a result of species sequence variation that was not reflected in the NCBI database. TaqMan

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

545 assays for *C. subsalsa* and *C. marina/ovata/antiqua* did not provide positive results with any of
546 the samples.
547

548 ***Silicoflagellates (Dictyochales)***

549 Silicoflagellates are a small group of algae, many of which possess a siliceous skeleton at certain
550 stages of their life cycle. Some species from this group cause fish kills (Henriksen et al., 1993),
551 including farmed salmon mortality events in BC (Haigh and Esenkulova, 2014; Haigh et al.,
552 2014, Haigh et al. 2019). All four amplicons underperformed in detecting algal species and
553 genera from the silicoflagellate group. Most of the samples (11 out of 15) where taxa from this
554 group were observed by microscopy did not have reads annotated as silicoflagellates. The
555 definitive reason for poor metabarcoding identification performance to this target remains
556 unknown, however it could be related to the extraction process, to the primers or amplification
557 process, or due to incomplete databases.
558

559 ***Silicoflagellates (Dictyochales): Dictyocha***

560 Skeleton-containing cells of *D. speculum* are very easily identified by microscopy based on their
561 unique shape and size. Blooms of this species have been linked with fish kills around the world
562 (*e.g.* Henriksen et al., 1993) including in BC (Haigh et al., 2019; Haigh and Esenkulova, 2014).
563 In our study, six field samples contained *D. speculum* cells, based on microscopy observations,
564 but there were no reads assigned to the *Dictyocha* species, genus, or family by any of the NGS
565 amplicons. One of these samples had a very low number of reads (*i.e.* 12 reads) detected by the
566 18S-diatom amplicon, but was assigned only to the order level. These results show that with the
567 current sequencing database, the NGS approach can miss targets at the species, genus, and even
568 family levels. Further work is needed to improve the sequencing database with sequences from
569 voucher specimens to determine which assays are capable of differentiating this genus, or species
570 within this genus, from other closely related taxa.
571

572 ***Silicoflagellates (Dictyochales): Pseudochattonella verruculosa***

573 Both species in the *Pseudochattonella* genus (*i.e. P. farcimen* and *P. verruculosa*) have been
574 implicated in wild and farmed fish kills (*e.g.* Jakobsen et al., 2012). In BC, *Pseudochattonella* sp.
575 blooms have been associated with farmed fish mortalities since 2007 (HAMP, unpublished data).

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

576 For this project, one culture from a fish-killing event (Haigh et al., 2014) was established and
577 five field samples were collected. High and comparable loads (>2,000 reads) of
578 *Pseudochattonella* genus were detected in the culture sample by 18S-diatom and 18S-
579 dinoflagellate amplicons. For the former, ~1% of the total reads were assigned to the species
580 level for *P. farcimen*. This culture sample was also the only one that provided qPCR
581 amplification results (Ct=14) with the *P. verruculosa* assay. A very low number of reads (<20
582 reads) of *Pseudochattonella* were detected in one of the field samples, where its presence was
583 indicated based on microscopy and by 18S-diatom and 18S-dinoflagellate runs, but not by qPCR.
584 This is the first published confirmation of both qPCR and metabarcoding identification of
585 *Pseudochattonella* linked to a fish-killing event in Canada.
586
587 ***Silicoflagellates (Dictyochales): Pseudopedinella***
588 Toxicity of *P. pyriformis* (previously known as *P. pyriforme*) was recently discovered in
589 laboratory studies (Skjelbred et al., 2011). Based on microscopic identification, this species has
590 been linked to farmed salmon mortality in BC (HAMP, unpublished data). In this work, one
591 culture and two field samples were available for molecular analysis. All these samples were
592 obtained from a fish-killing event on the west coast of Vancouver Island. To our knowledge, this
593 is the first published report of ichthyotoxic effects of *Pseudopedinella* species in the field.
594 Metabarcoding did not detect *Pseudopedinella* in the available samples, and it did not identify
595 any reads from the silicoflagellates taxon. The majority of reads in the culture of suspected
596 *Pseudopedinella* sp. was the diatom *Plagiostriata goreensis* by the 18S-diatom amplicon (98%)
597 and "not assigned" in the 18S-dinoflagellate and LSU amplicons (94%, 99%, respectively).
598 While there was therefore no molecular confirmation of the microscopic identification of
599 *Pseudopedinella* in the fish-kill-related sample, the established culture was certainly not a diatom
600 and contained *Pseudopedinella*-like cells (small flagellates). To improve this detection method, a
601 positive control would need to be included, such as a purchased *Pseudopedinella* culture, in
602 order to ensure that the selected amplicons can identify this genus. However, given the putative
603 misidentification of the culture of suspected *Pseudopedinella*, it is also possible that this is a
604 database issue.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

605 ***Haptophytes***

606 Toxic haptophytes in the *Prymnesium* and *Chrysochromulina* genera are known to cause fish
607 mortality (Moestrup, 2003). In our work, the best results were obtained using the 18S-
608 dinoflagellate and LSU amplicons.
609

610 ***Haptophytes: Chrysochromulina***

611 Haptophytes may be identified with light microscopy to the genus level, but this is challenging
612 due to small cell size, and species level identification is only possible with scanning electron
613 microscopy. In BC, blooms of suspected *Chrysochromulina* spp. have been linked to fish kills on
614 salmon farms since 2000 (HAMP, unpublished data). Our study had four field samples where
615 *Chrysochromulina* presence was suspected based on microscopy, but two of these samples had
616 low DNA yields and resulted in <200 total reads per sample. The 18S-dinoflagellate amplicon
617 detected *Haptolina fragaria* (>100 reads) in the other two samples. *Haptolina fragaria*
618 (presumably a non-toxic Prymnesiales species), originally described as *Chrysochromulina*
619 *fragaria* (Eikrem and Edvardsen, 1999) but later described to a new genus from ribosomal DNA
620 phylogenetics (Edvardsen et al., 2011). Low levels of *Chrysochromulina* spp. were also detected
621 by the LSU amplicon (220 reads, 0.35%) in one field sample where its presence was not
622 indicated based on microscopy. With the importance of haptophyte species in fish-kills, and the
623 difficulty in microscopic identification of these species, this is an area where better detection by
624 molecular methods would be of strong practical use.
625

626 ***Cyanobacteria***

627 Cyanobacteria, or blue-green algae, are highly diverse aquatic bacteria with over 2,000 species
628 (Nabout et al., 2013). More than 55 of them have been shown to produce toxins that are harmful
629 to humans, as well as other terrestrial and aquatic life (Cronberg, 2003). Due to their generally
630 small size and limited significance to marine finfish, cyanobacteria are not recorded and
631 identified during routine harmful algae analysis in BC by HAMP, unless they appear to dominate
632 in the sample. Here, only one sample (mixed culture) had evidence of possible cyanobacteria
633 presence (based on microscopy), and was mixed with green algae. The cyanobacteria presence in
634 this sample was not confirmed by molecular techniques, but cyanobacteria were detected in other
635 samples by molecular techniques where its presence was not described based on microscopy.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

636 Nine samples provided amplification results with qPCR (Ct=17–29), but only one of these
637 samples had evidence for cyanobacteria presence based on metabarcoding, using the 16S
638 amplicon (Synechococcaceae, 101 reads, 0.1%). Whether the qPCR results are species specific,
639 or whether the 16S metabarcoding was missing these detections remains unknown, and requires
640 further study.
641

642 ***Summary of Methods Evaluation***

643 Development of molecular methods for harmful taxa detection and abundance estimation is an
644 active area of research. The present study allowed for an evaluation and cross validation of
645 traditional microscopy and current molecular techniques and assays for each taxon relevant to
646 BC HABs (Table 7). The best molecular technique, as determined by highest detection rate using
647 currently available assays was identified based on the results such as taxon resolution, the
648 conclusiveness of the detection, and whether it was confirmatory of microscopy results.
649 Molecular techniques, and metabarcoding in particular, proved to be a very promising
650 complement, and in some cases potentially an alternative given additional benchmarking, to
651 standard microscopy. However, there is still much work to be done to develop a curated database
652 with voucher specimens to ensure adequate representation and detectability of all of the
653 important HABs forming species. The comprehensive nature of the metabarcoding approach is
654 another benefit, where many species can be simultaneously detected even from different
655 taxonomic levels. Information obtained during this research provides a foundation to build upon,
656 for example, by using the sequencing results to identify species presence, to develop specific
657 qPCR primers, and to identify existing gaps in sequence databases.

658 Quantitative PCR is a cost effective approach and can be used for presence/absence as
659 well as quantitative estimates of targeted species, however it requires existing knowledge of
660 potential HABs taxa that may be present in monitored samples. In our study, existing TaqMan
661 assays were found for only four out of 14 targeted taxa. Assays for detection of *A. tamarense*
662 (A.tam1), *H. akashiwo* (H.aka1), and *P. verruculosa* (P.ver1) performed well and could be
663 implemented in routine monitoring. Ten taxa that were identified by microscopy did not have a
664 molecular assay available during the time of this study, including *Chrysochromulina* spp.,
665 *Cochlodinium fulvescens*, *C. concavicornis*, *C. convolutus, Dictyocha* spp., *Dinophysis* spp., *K.*
666 *mikimotoi*, *Protoceratium reticulatum*, and *Pseudo-nitzschia* spp. These particular taxa are an

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

667 important priority for assay development for application in BC. Sequencing data, acquired during
668 this work, provides an ability to develop tailored qPCR assays.

669 The metabarcoding approach offered accurate identification of multiple species and was
670 superior when applied to the detection and identification of cryptic species that cannot be
671 differentiated to the species level by light microscopy (*e.g. Chrysochromulina*, *Prymnesium*). It
672 generated high-throughput data, providing community composition diversity information for the
673 samples, including capturing some potentially harmful taxa that were not originally targeted, as
674 well as providing some cursory information on relative abundance, although the reliability of
675 abundance estimates has not yet been conducted. The presence of a large number of unannotated
676 reads in a number of samples suggests that the sequences of a number of species are still missing
677 from the public databases. This will improve with curated databases, which will be generated as
678 this method and application continues to be developed.

679 The highly sensitive nature of this technique emphasised the need for extremely careful
680 sample handling, as detections of pure cultures transferring small proportions of taxa between
681 samples was detected, probably caused during subsampling of cultures. Since the laboratory
682 component of this study, the increased use of metabarcoding in many areas of research (*e.g.*
683 environmental DNA, or eDNA) applications, has resulted in specialized sample collection,
684 processing workflows, laboratory environments, and establishing threshold detections designed
685 to reduce contamination risks.

686 Although metabarcoding processing time and costs may limit uptake currently, with
687 increased development it is highly likely to become a useful technique for analysing samples
688 collected during active HABs, leading to an unprecedented monitoring opportunity of HABs in
689 terms of sensitivity and precision.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

690
CONCLUSIONS

691 Applying light microscopy and various molecular techniques to culture and field samples
692 containing multiple harmful algal species allowed cross validation of these techniques and
693 offered a significant foundation for choosing appropriate techniques for targeted taxa in the
694 future. DNA yields were considerably higher for the cultures and filtered field samples, whereas
695 pelleted samples were often unusable. While TaqMan assays were available only for four out of
696 14 HAB taxa of concern in BC, assays for detection of *A. tamarense*, *H. akashiwo*, and *P.*
697 *verruculosa* provided adequate identification results. This indicates a need for the development
698 of primers and probes for the rest of the harmful species to allow cost effective detection of
699 many species simultaneously. Sequencing data obtained during this study enabled the
700 development of new qPCR assays tailored for species within the northeastern Pacific Ocean.

701 Metabarcoding with a combination of markers (*i.e.* 16S, 18S-diatom, 18S-dinoflagellate,
702 and LSU) allowed the identification of over 350 taxa and proved to be an unmatched technique
703 for phytoplankton community structure analysis. Different markers had different strengths for
704 particular taxa, although result congruence was observed among amplicons. The 18S-diatom
705 amplicon identified harmful taxa from the diatom and raphidophyte groups, and the 18S-
706 dinoflagellate and LSU amplicons, producing similar results to each other, overall provided the
707 best identification for harmful algae from dinoflagellate and haptophytes groups, as well as the
708 raphidophyte *H. akashiwo*. Although cyanobacteria were detected by only the 16S amplicon, this
709 amplicon was weak for most of the HABs species. All amplicons appeared to underperform for
710 identification of silicoflagellate algae, although the reason for this remains unknown. The
711 isolated culture from the fish-killing event associated with a *P. verruculosa* bloom (morphology-
712 based identification) was confirmed as *P. verruculosa* by qPCR and *Pseudochattonella* spp. by
713 NGS. This is the first record of both PCR and metabarcoding confirmation of *Pseudochattonella*
714 associated with a fish kill in Canada. Overall, the combination of morphology and molecular-
715 based identification, if implemented, will greatly improve HABs monitoring, help mitigate issues
716 caused by HABs, and aid in better understanding the dynamics of the phenomenon. This work
717 demonstrates a pressing need to tailor qPCR assays, to improve reference databases, and apply a
718 multiple marker approach for metabarcoding of diverse taxa.

719

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

720 **DATA AVAILABILITY**

721 Analysis pipeline: https://github.com/bensutherland/eDNA_metabarcoding

722 Raw data: raw sequence was uploaded to SRA under BioProject PRJNA544881 within
723 BioSample accessions SAMN11865982-SAMN11866125.

724

725 **SUPPLEMENTAL INFORMATION**

726 **Supplemental File S1.** Number of reads for each taxonomic group identified with the 16S
727 amplicon.

728 **Supplemental File S2.** Number of reads for each taxonomic group identified with the 18S
729 diatom amplicon.

730 **Supplemental File S3.** Number of reads for each taxonomic group identified with the 18S
731 dinoflagellate amplicon.

732 **Supplemental File S4.** Number of reads for each taxonomic group identified with the LSU
733 amplicon.

734 **Supplemental File S5.** Average number of algae taxa (t) and algae species (s) detected by the
735 different amplicons.

736 **Supplemental File S6.** Taxa detected by microscopy, the closest match as detected by NGS, the
737 percentage of total reads for that sample represented by the match, and the percentage of
738 unknown reads in the sample.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

739
ACKNOWLEDGEMENTS

740 This project was supported by the Aquaculture Collaborative Research and Development
741 Program (ACRDP; Project number: P-12-01-003) of Fisheries and Oceans Canada and the
742 British Columbia aquaculture industry including Creative Salmon Company Ltd., Grieg Seafood
743 BC Ltd., Mainstream Canada (now Cermaq Canada), Marine Harvest Canada Inc. (now Mowi
744 Canada West), Cleanwater Shellfish Ltd., Island Scallops Ltd., Little Wing Oysters Ltd., Mac's
745 Oysters Ltd., Nelson Island Sea Farms Ltd., and Taylor Shellfish Canada ULC. Yves Perreault of
746 Little Wing Oysters Ltd. and Dave Guhl of Marine Harvest Canada Inc. are acknowledged for
747 helping with sediment sampling at Okeover Inlet and Quatsino Sound, respectively. Crew of the
748 Deep Bay Marine Station enabled sediment sampling in Baynes Sound. Laurie Keddy (Pacific
749 Biological Station, DFO) is thanked for maintaining established cultures and providing advice on
750 isolation and culturing.

751

752
AUTHOR CONTRIBUTIONS

753 Wrote the proposal to obtain funding: KM, CP, NH. Conceived and designed the experiments:
754 NH, SE, KM, AT, CP. Performed the experiments: SE, NH, AT. Bioinformatics: BS. Analyzed
755 the data: SE, AT, BS, NH. Contributed reagents/materials/analysis tools: KM, CP, NH. Wrote
756 the manuscript: SE, BS, AT. Revised and approved the manuscript: SE, BS, AT, KM, CP, NH.

757

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

758 # REFERENCES

759 Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J.,
760 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search
761 programs. Nucleic Acids Res. https://doi.org/10.1093/nar/25.17.3389

762 Andersen, N., 2012. Ichthyotoxic algae and their effects on fish. … Univ. Det Natur-og ….

763 Andersen, P., Throndsen, J., 2004. Estimating cell numbers, in: Manual on Harmful Marine
764 Microalgae. Monographs on Oceanographic Methodology No. 11.

765 Anderson, D.M., Fukuyo, Y., Matsuoka, K., 2003. Cyst methodologies. Man. Harmful Mar.
766 Microalgae.

767 Andrews, S., Babraham Bioinformatics, 2010. FastQC: A quality control tool for high
768 throughput sequence data. Manual. https://doi.org/citeulike-article-id:11583827

769 Antonella, P., Luca, G., 2013. The quantitative real-time PCR applications in the monitoring of
770 marine harmful algal bloom (HAB) species. Environ. Sci. Pollut. Res.
771 https://doi.org/10.1007/s11356-012-1377-z

772 Balech, E., 1985. The genus Alexandrium or Gonyaulax of the Tamarensis group, in:
773 Proceedings of the Third International Conference on Toxic Dinoflagellates, St. Andrews,
774 New Brunswick, Canada, June 8-12, 1985.

775 Bowers, H.A., Tengs, T., Glasgow, J., Burkholder, J.M., Rublee, P.A., Oldach, D.W., 2000.
776 Development of real-time PCR assays for rapid detection of Pfiesteria piscicida and related
777 dinoflagellates. Appl. Environ. Microbiol. https://doi.org/10.1128/AEM.66.11.4641-
778 4648.2000

779 Bowers, H.A., Tomas, C., Tengs, T., Kempton, J.W., Lewitus, A.J., Oldach, D.W., 2006.
780 Raphidophyceae [chadefaud ex silva] systematics and rapid identification: Sequence
781 analyses and real-time PCR assays. J. Phycol. https://doi.org/10.1111/j.1529-
782 8817.2006.00285.x

783 Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., Coissac, E., 2016. obitools: a unix-
784 inspired software package for DNA metabarcoding. Mol. Ecol. Resour.
785 https://doi.org/10.1111/1755-0998.12428

786 Cembella, A.D., 2003. Chemical ecology of eukaryotic microalgae in marine ecosystems.
787 Phycologia 42, 420–447. https://doi.org/10.2216/i0031-8884-42-4-420.1

788 Chamberlain, S.A., Szöcs, E., 2013. taxize: taxonomic search and retrieval in R. F1000Research.
789 https://doi.org/10.12688/f1000research.2-191.v2

790 Coyne, K.J., Handy, S.M., Demir, E., Whereat, E.B., Hutchins, D.A., Portune, K.J., Doblin,
791 M.A., Cary Craig, S., 2005. Improved quantitative real-time PCR assays for enumeration of
792 harmful algal species in field samples using an exogenous DNA reference standard. Limnol.
793 Oceanogr. Methods. https://doi.org/10.4319/lom.2005.3.381

794 Delaney, J.A., Ulrich, R.M., Paul, J.H., 2011. Detection of the toxic marine diatom Pseudo-
795 nitzschia multiseries using the RuBisCO small subunit (rbcS) gene in two real-time RNA
796 amplification formats. Harmful Algae. https://doi.org/10.1016/j.hal.2011.07.005

797 Department of Fisheries and Oceans Canada [WWW Document], n.d.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

798 Eckford-Soper, L.K., Daugbjerg, N., 2015. Examination of six commonly used laboratory
799 fixatives in HAB monitoring programs for their use in quantitative PCR based on Taqman
800 probe technology. Harmful Algae. https://doi.org/10.1016/j.hal.2014.12.007

801 Edvardsen, B., Eikrem, W., Throndsen, J., Sáez, A.G., Probert, I., Medlin, L.K., 2011.
802 Ribosomal DNA phylogenies and a morphological revision provide the basis for a revised
803 taxonomy of the Prymnesiales (haptophyta). Eur. J. Phycol. 46, 202–228.
804 https://doi.org/10.1080/09670262.2011.594095

805 Eikrem, W., Edvardsen, B., 1999. Chrysochromulina fragaria sp nov (Prymnesiophyceae), a new
806 haptophyte flagellate from Norwegian waters. Phycologia 38, 149–155. https://doi.org/DOI
807 10.2216/i0031-8884-38-2-149.1

808 Eiler, A., Drakare, S., Bertilsson, S., Pernthaler, J., Peura, S., Rofner, C., Simek, K., Yang, Y.,
809 Znachor, P., Lindström, E.S., 2013. Unveiling Distribution Patterns of Freshwater
810 Phytoplankton by a Next Generation Sequencing Based Approach. PLoS One.
811 https://doi.org/10.1371/journal.pone.0053516

812 Esenkulova, S., Haigh, N., 2012. First report of Dinophysis species causing diarrhetic shellfish
813 poisoning in British Columbia, Canada. Harmful Algae News UNESCO 45, 16–17.

814 Esenkulova, S., Luinenburg, O., Neville, C.M., Trudel, M., 2015. Observations of Heterosigma
815 akashiwo bloom and associated wild salmon lethargic behaviour in Cowichan Bay, Canada.
816 Harmful Algae News 50, 16–18.

817 Ewels, P., Magnusson, M., Lundin, S., Käller, M., 2016. MultiQC: Summarize analysis results
818 for multiple tools and samples in a single report. Bioinformatics.
819 https://doi.org/10.1093/bioinformatics/btw354

820 G. Cronberg, E.J.C. and W.W.C., 2003. Taxonomy of harmful cyanobacteria, in: Manual on
821 Harmful Marine Microalgae. https://doi.org/10.1016/B978-0-12-391499-6.00008-6

822 Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: Ten years of next-
823 generation sequencing technologies. Nat. Rev. Genet. https://doi.org/10.1038/nrg.2016.49

824 Gray, M., Wawrik, B., Paul, J., Casper, E., 2003. Molecular detection and quantitation of the red
825 tide dinoflagellate Karenia brevis in the marine environment. Appl. Environ. Microbiol.
826 https://doi.org/10.1128/AEM.69.9.5726-5730.2003

827 Guillard, 1978. . Counting slides. In: Sournia, A. (Ed.), Phytoplankton Manual. Monographs on
828 Oceanographic Methodology 6, UNESCO user manual. https://doi.org/10.2216/i0031-8884-
829 19-4-341.1

830 Haigh, N., Brown, T., Johnson, D., 2019. Ichthyotoxic skeleton-forming silicoflagellates in
831 British Columbia, Canada; results from the Harmful Algae Monitoring Program, 1999 –
832 2017, in: Hess, P. (Ed.), . Proceedings of the 18th International Conference on Harmful
833 Algae.

834 Haigh, N., Esenkulova, S., 2014. Economic losses to the British Columbia salmon aquaculture
835 industry due to harmful algal blooms, 2009-2012. PICES Sci. Rep. 2.

836 Haigh, N., Esenkulova, S., Pudota, J., Pearce, C.M., Keddy, L.J., Tabata, A., Miller, K.M., 2014.
837 First confirmed report of fish-killing Pseudochattonella species (Dictyochophyceae) on the
838 west coast of Canada, in: MacKenzie, L. (Ed.), Proceedings of the 16th International
839 Conference on Harmful Algae. Cawthron Institute, Nelson, pp. 270–273.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

840 Hallegraeff, G.M., 2004. Harmful algal blooms: a global overview, Manual on Harmful Marine
841 Microalgae. https://doi.org/10.1016/B978-0-12-391499-6.00008-6

842 Handy, S.M., Demir, E., Hutchins, D.A., Portune, K.J., Whereat, E.B., Hare, C.E., Rose, J.M.,
843 Warner, M., Farestad, M., Cary, S.C., Coyne, K.J., 2008. Using quantitative real-time PCR
844 to study competition and community dynamics among Delaware Inland Bays harmful algae
845 in field and laboratory studies. Harmful Algae. https://doi.org/10.1016/j.hal.2007.12.018

846 Handy, S.M., Hutchins, D.A., Cary, S.C., Coyne, K.J., 2006. Simultaneous enumeration of
847 multiple raphidophyte species by quantitative real-time PCR: Capabilities and limitations.
848 Limnol. Oceanogr. Methods. https://doi.org/10.4319/lom.2006.4.193

849 Harrison, P.J., Waters, R.E., Taylor, F.J.R., 1980. A BROAD SPECTRUM ARTIFICIAL SEA
850 WATER MEDIUM FOR COASTAL AND OPEN OCEAN PHYTOPLANKTON. J.
851 Phycol. https://doi.org/10.1111/j.0022-3646.1980.00028.x

852 Hasle, G.R., 1978. Identification problems. General recommendations. Phytoplankt. manual.
853 UNESCO Monogr. Ocean. Method 6, 125–128.

854 Hasle, G.R., Syvertesen, E.E., 1996. Marine diatoms, in: Identifying Marine Diatoms and
855 Dinoflagellates. p. 598. https://doi.org/DOI: 10.1016/B978-012693015-3/50005-X

856 Henriksen, P., Knipschildt, F., Moestrup, Ø., Thomsen, H. a., 1993. Autecology, life history and
857 toxicology of the silicoflagellate Dictyocha speculum (Silicoflagellata, Dictyochophyceae).
858 Phycologia 32, 29–39. https://doi.org/10.2216/i0031-8884-32-1-29.1

859 Horner, R.A., Garrison, D.L., Plumley, F.G., 1997. Harmful algal blooms and red tide problems
860 on the U.S. west coast. Limnol. Oceanogr. https://doi.org/10.4319/lo.1997.42.5_part_2.1076

861 Hosoi-Tanabe, S., Sako, Y., 2005. Species-specific detection and quantification of toxic marine
862 dinoflagellates Alexandrium tamarense and A. catenella by real-time PCR assay. Mar.
863 Biotechnol. https://doi.org/10.1007/s10126-004-4128-4

864 Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.J., Tappu,
865 R., 2016. MEGAN Community Edition - Interactive Exploration and Analysis of Large-
866 Scale Microbiome Sequencing Data. PLoS Comput. Biol.
867 https://doi.org/10.1371/journal.pcbi.1004957

868 Jakobsen, R., Hansen, P.J., Daugbjerg, N., Andersen, N.G., 2012. The fish-killing
869 dictyochophyte Pseudochattonella farcimen: Adaptations leading to bloom formation during
870 early spring in Scandinavian waters. Harmful Algae 18, 84–95.
871 https://doi.org/10.1016/j.hal.2012.04.008

872 John, U., Litaker, R.W., Montresor, M., Murray, S., Brosnahan, M.L., Anderson, D.M., 2014.
873 Formal revision of the alexandrium tamarense species complex (dinophyceae) taxonomy:
874 The introduction of five species with emphasis on molecular-based (rDNA) classification.
875 Protist. https://doi.org/10.1016/j.protis.2014.10.001

876 Kamikawa, R., Asai, J., Miyahara, T., Murata, K., Oyama, K., Yoshimatsu, S., Yoshida, T.,
877 Sako, Y., 2006. Application of a Real-time PCR Assay to a Comprehensive Method of
878 Monitoring Harmful Algae. Microbes Environ. https://doi.org/10.1264/jsme2.21.163

879 Kudela, R.M., Gobler, C.J., 2012. Harmful dinoflagellate blooms caused by Cochlodinium sp.:
880 Global expansion and ecological strategies facilitating bloom formation. Harmful Algae 14,
881 71–86. https://doi.org/10.1016/j.hal.2011.10.015

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

882 Lilly, E.L., Halanych, K.M., Anderson, D.M., 2007. Species boundaries and global biogeography
883 of the Alexandrium tamarense complex (Dinophyceae). J. Phycol.
884 https://doi.org/10.1111/j.1529-8817.2007.00420.x

885 Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time
886 quantitative PCR and the 2-ΔΔCT method. Methods.
887 https://doi.org/10.1006/meth.2001.1262

888 Manning, S.R., La Claire, J.W., 2010. Prymnesins: Toxic metabolites of the golden alga,
889 Prymnesium parvum Carter (Haptophyta). Mar. Drugs. https://doi.org/10.3390/md8030678

890 Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
891 EMBnet.journal. https://doi.org/10.14806/ej.17.1.200

892 Medlin, L.K., Orozco, J., 2017. Molecular techniques for the detection of organisms in aquatic
893 environments, with emphasis on harmful algal bloom species. Sensors (Switzerland).
894 https://doi.org/10.3390/s17051184

895 Miller, K.M., Gardner, I.A., Vanderstichel, R., Brunley, T., Schulze, A.D., Li, K.H., Kaukinen,
896 K.H., Ming, T.J., Ginther, N.G., 2016. Report on the performance evaluation of the
897 Fluidigm BioMark platform for high-throughput microbe monitory in salmon. DFO Can.
898 Sci. Advis. Sec. Res. Doc.

899 Moestrup, Ø., 2003. Taxonomy of toxic haptophytes (prymnesiophytes), in: Manual on Harmful
900 Marine Microalgae. pp. 433–463.

901 Moestrup, Ø., Codd, G.A., Elbrächter, M., Faust, M.A., Fraga, S., Fukuyo, Y., Cronberg, G.,
902 Halim, Y. (Eds), 2008. IOC-UNESCO Taxonomic Reference List of Harmful Micro Algae.
903 Available online at http://www.marinespecies.org/HAB.

904 Nabout, J.C., da Silva Rocha, B., Carneiro, F.M., Sant'Anna, C.L., 2013. How many species of
905 Cyanobacteria are there? Using a discovery curve to predict the species number. Biodivers.
906 Conserv. 22, 2907–2918. https://doi.org/10.1007/s10531-013-0561-x

907 Ondov, B.D., Bergman, N.H., Phillippy, A.M., 2011. Interactive metagenomic visualization in a
908 Web browser. BMC Bioinformatics. https://doi.org/10.1186/1471-2105-12-385

909 Park, T.-G., Park, Y.-T., Lee, Y., 2009. Development of a SYTO9 based real-time PCR probe
910 for detection and quantification of toxic dinoflagellate Karlodinium veneficum
911 (Dinophyceae) in environmental samples . Phycologia. https://doi.org/10.2216/08-52.1

912 Park, T.G., Park, G.H., Park, Y.T., Kang, Y.S., Bae, H.M., Kim, C.H., Jeong, H.J., Lee, Y.,
913 2009. Identification of the dinoflagellate community during Cochlodinium polykrikoides
914 (Dinophyceae) blooms using amplified rDNA melting curve analysis and real-time PCR
915 probes. Harmful Algae. https://doi.org/10.1016/j.hal.2008.09.003

916 Place, A.R., Bowers, H.A., Bachvaroff, T.R., Adolf, J.E., Deeds, J.R., Sheng, J., 2012.
917 Karlodinium veneficum-The little dinoflagellate with a big bite. Harmful Algae 14, 179–
918 195. https://doi.org/10.1016/j.hal.2011.10.021

919 R Core Team. [WWW Document], 2018. . R A Lang. Environ. Stat. Comput. R Found. Stat.
920 Comput. Vienna. URL https://www.r-project.org

921 Rensel, J.E., Whyte, J.N.C., 2004. Finfish mariculture and harmful algal blooms. Man. harmful
922 Mar. microalgae. Monogr. Oceanogr. Methodol. 11, 693–722.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

923 Rinta-Kanto, J.M., Ouellette, A.J.A., Boyer, G.L., Twiss, M.R., Bridgeman, T.B., Wilhelm,
924 S.W., 2005. Quantification of toxic Microcystis spp. during the 2003 and 2004 blooms in
925 western Lake Erie using quantitative real-time PCR. Environ. Sci. Technol.
926 https://doi.org/10.1021/es048249u

927 Satake, M., MacKenzie, L., Yasumoto, T., 1998. Identification of Protoceratium reticulatum as
928 the biogenetic origin of yessotoxin. Nat. Toxins 5, 164–167.
929 https://doi.org/10.1002/19970504NT7

930 Scholin, C., Doucette, G., Jensen, S., Roman, B., Pargett, D., Marin III, R., Preston, C., Jones,
931 W., Feldman, J., Everlove, C., Harris, A., Alvarado, N., Massion, E., Birch, J., Greenfield,
932 D., Wheeler, K., Vrijenhoek, R., Mikulski, C., Jones, K., 2011. Remote Detection of Marine
933 Microbes, Small Invertebrates, Harmful Algae, and Biotoxins using the Environmental
934 Sample Processor (ESP). Oceanography. https://doi.org/10.5670/oceanog.2009.46

935 Scott, A., Szoecs, E., Foster, Z., Ram, K., Baumgartner, J., Donnell, J.O., 2019. Package ‘ taxize
936 .’

937 Skjelbred, B., Horsberg, T.E., Tollefsen, K.E., Andersen, T., Edvardsen, B., 2011. Toxicity of
938 the ichthyotoxic marine flagellate Pseudochattonella (Dictyochophyceae, Heterokonta)
939 assessed by six bioassays. Harmful Algae 10, 144–154.
940 https://doi.org/10.1016/j.hal.2010.08.007

941 Tange, O., 2011. GNU Parallel□: The Command-Line Power Tool. USENIX□;login;

942 Taylor, F.J.R., Harrison, P.J., 2002. Harmful algal blooms in western Canadian coastal waters.
943 Harmful algal Bloom. PICES Reg. North Pacific. PICES Sci. Rep. 1965, 77–88.
944 https://doi.org/10.4103/1735-5362.213981

945 Taylor, F.J.R., Hoppenrath, M., Saldarriaga, J.F., 2008. Dinoflagellate diversity and distribution.
946 Biodivers. Conserv. https://doi.org/10.1007/s10531-007-9258-3

947 Toebe, K., Alpermann, T.J., Tillmann, U., Krock, B., Cembella, A., John, U., 2013. Molecular
948 discrimination of toxic and non-toxic Alexandrium species (Dinophyta) in natural
949 phytoplankton assemblages from the Scottish coast of the North Sea. Eur. J. Phycol.
950 https://doi.org/10.1080/09670262.2012.752870

951 Toebe, K., Joshi, A.R., Messtorff, P., Tillmann, U., Cembella, A., John, U., 2012. Molecular
952 discrimination of taxa within the dinoflagellate genus Azadinium, the source of azaspiracid
953 toxins. J. Plankton Res. https://doi.org/10.1093/plankt/fbs077

954 Toyoda, K., Nagasaki, K., Tomaru, Y., 2010. Application of real-time PCR assay for detection
955 and quantification of bloom-forming diatom Chaetoceros tenuissimus Meunier. Plankt.
956 Benthos Res. https://doi.org/10.3800/pbr.5.56

957 Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P.F., Bellemain, E.,
958 Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp,
959 G.H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.-M., Peroux, T., Crivelli, A.J., Olivier,
960 A., Acqueberge, M., Le Brun, M., Møller, P.R., Willerslev, E., Dejean, T., 2016. Next-
961 generation monitoring of aquatic biodiversity using environmental DNA metabarcoding.
962 Mol. Ecol. https://doi.org/10.1111/mec.13428

963 Vancouver, G., 1798. Seaman poisoned by mussels, in: Robinson (Ed.), A Voyage of Discovery
964 to the North Pacific Ocean and Round the World. Paternoster-Row and J. Edwards, London.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

965 Venrick, E.L., 1978. Systematic sampling in a planktonic ecosystem. Fish. Bull. 76, 617–627.

966 Wall, D., Dale, B., 1967. The resting cysts of modern marine dinoflagellates and their
967 palaeontological significance. Rev. Palaeobot. Palynol. https://doi.org/10.1016/0034-
968 6667(67)90165-0

969 Whyte, J.N.C. (Ian), Haigh, N., Ginther, N.G., Keddy, L.J., 2001. First record of blooms of
970 Cochlodinium sp. (Gymnodiniales, Dinophyceae) causing mortality to aquacultured salmon
971 on the west coast of Canada. Phycologia 40, 298–304. https://doi.org/10.2216/i0031-8884-
972 40-3-298.1

973 Whyte, J.N.C., Davis, J.C., Forbes, J.R., 1997. Harmful algae in Canadian waters and
974 management strategies. Oceanogr. Lit. Rev. 7, 1234–1235.

975 Wickham, H., 2016. ggplot2: elegant graphics for data analysis. Springer.

976 Yuan, J., Mi, T., Zhen, Y., Yu, Z., 2012. Development of a rapid detection and quantification
977 method of Karenia mikimotoi by real-time quantitative PCR. Harmful Algae.
978 https://doi.org/10.1016/j.hal.2012.03.004

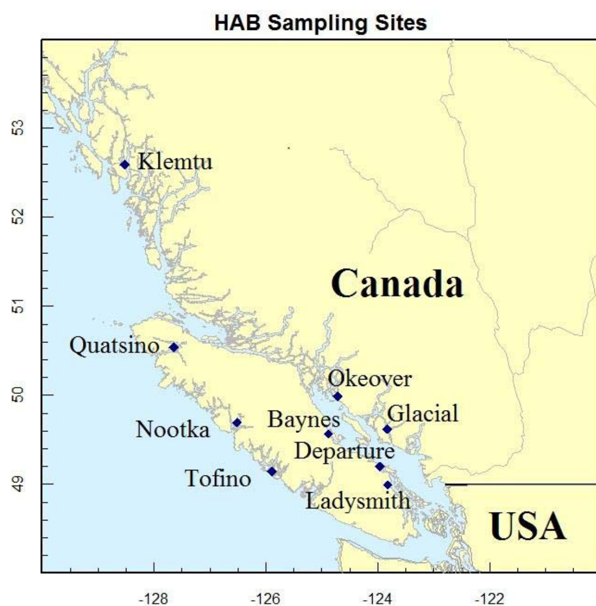
979 Zamor, R.M., Glenn, K.L., Hambright, K.D., 2012. Incorporating molecular tools into routine
980 HAB monitoring programs: Using qPCR to track invasive Prymnesium. Harmful Algae.
981 https://doi.org/10.1016/j.hal.2011.10.028

982

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

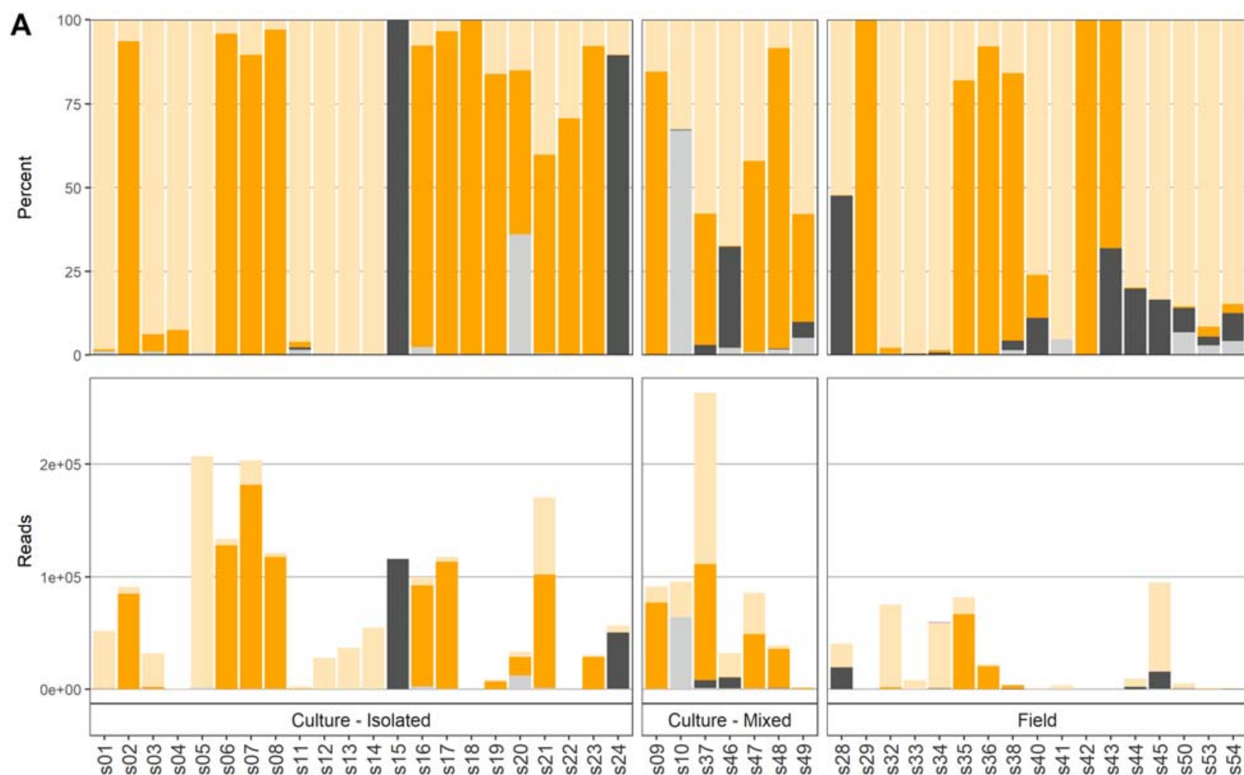
983 # FIGURES AND TABLES

984

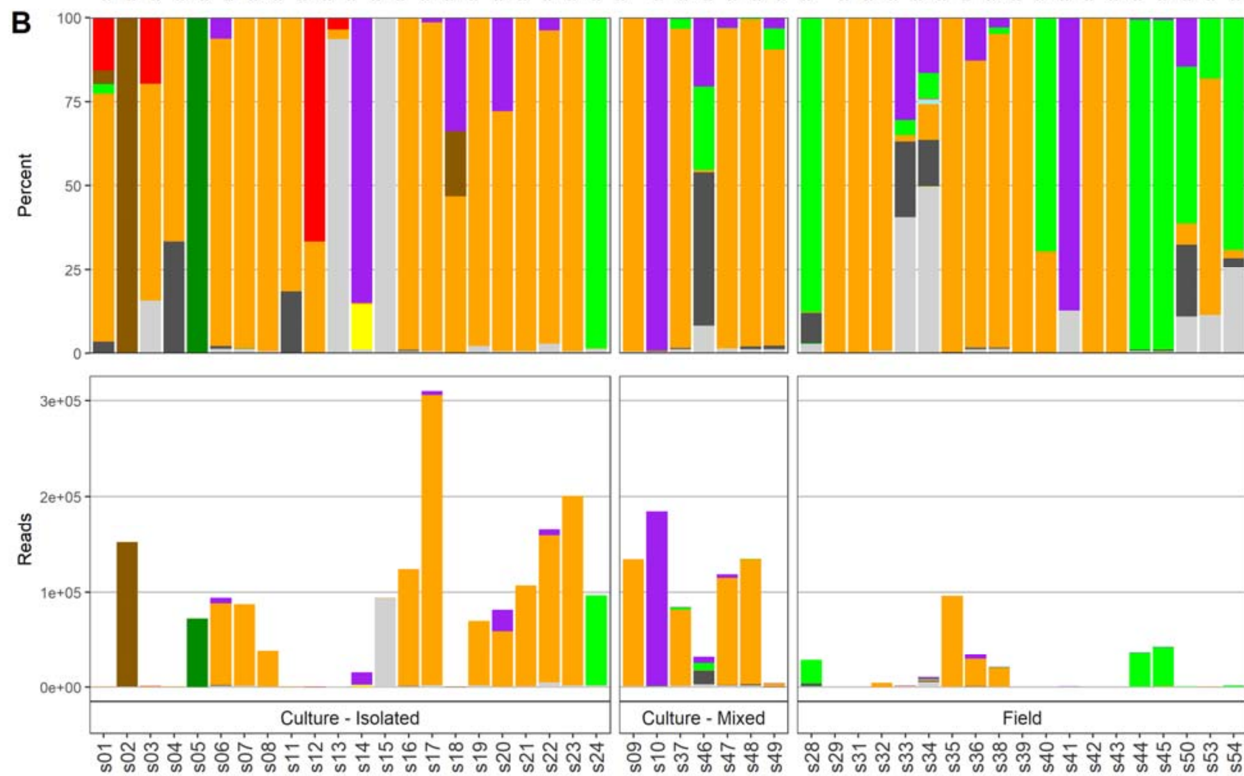


986 **Figure 1.** Map of Vancouver Island and British Columbia coast with locations of the sampling
987 stations.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



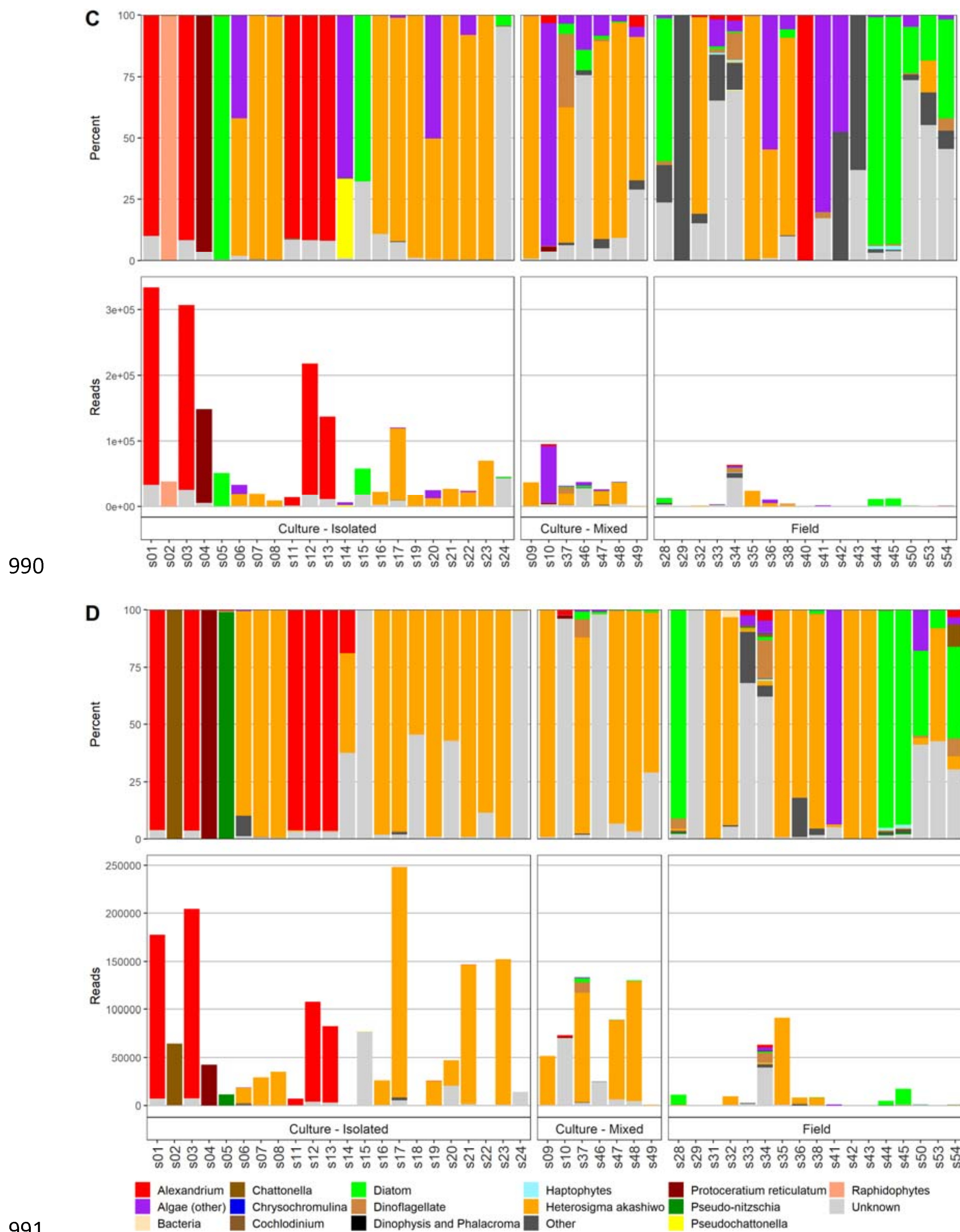
988



989

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

990



991

992 **Figure 2**. Percentages and number of reads per sample with emphasis on harmful and potentially
993 harmful algal species: (A) 16S, (B) 18S-Diatom, (C) 18S-Dinoflagellate, and (D) LSU.
994

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

995 **Table 1.** Sample identifiers, the microscopic observations, and their sampling origin and
996 collection location. Cultures from USA are denoted by an asterisk, with s01 from North Pacific
997 Ocean, Sequim Bay, Washington and s02 from North Atlantic Ocean, Indian River Bay,
998 Delaware.

ID	Microscopy observations	Sample origin	Sampling location
s01	*Alexandrium* cf *catenella*	PC	Sequim Bay*
s02	*Chattonella subsalsa*	PC	Indian River*
s03	*Alexandrium* sp.	IC	Okeover Inlet
s04	*Protoceratium reticulatum*	IC	Okeover Inlet
s05	*Pseudo-nitzschia* spp.	IC	Quatsino Sound
s06	*Heterosigma akashiwo*	IC	Quatsino Sound
s07	*Heterosigma akashiwo*	IC	Quatsino Sound
s08	*Heterosigma akashiwo*	IC	Quatsino Sound
s09	*Heterosigma akashiwo,* Raphidophyte (*Chattonella* sp.?), nanoflagellates	MC	Quatsino Sound
s10	*Alexandrium* sp., flagellates, green algae	MC	Okeover Inlet
s11	*Alexandrium* sp.	IC	Okeover Inlet
s12	*Alexandrium tamarense*	IC	Okeover Inlet
s13	*Alexandrium* sp.	IC	Okeover Inlet
s14	*Pseudochattonella* cf. *verruculosa*	IC	Quatsino Sound
s15	*Chaetoceros eibenii*	IC	Departure Bay
s16	*Heterosigma akashiwo*	IC	Quatsino Sound
s17	*Heterosigma akashiwo, Chattonella* sp.?	IC	Quatsino Sound
s18	*Heterosigma akashiwo*	IC	Quatsino Sound
s19	*Heterosigma akashiwo*	IC	Quatsino Sound
s20	*Heterosigma akashiwo*	IC	Ladysmith
s21	*Heterosigma akashiwo*	IC	Departure Bay
s22	*Heterosigma akashiwo*	IC	Klemtu
s23	*Heterosigma akashiwo*	IC	Glacial
s24	*Pseudopedinella* sp.	IC	Klemtu
s25	*Pseudopedinella* sp.	–	Quatsino Sound
s26	*Pseudopedinella* sp.	PWS	Quatsino Sound
s27	*Pseudopedinella* sp.	PWS	Quatsino Sound
s28	*Pseudopedinella* sp.	PWS	Quatsino Sound
s29	*Chattonella* sp., *Dictyocha* spp., *Heterosigma akashiwo, Prorocentrum gracile*	PWS	Nootka Sound
s30	*Chattonella* sp., *Dictyocha* spp., *Heterosigma akashiwo, Prorocentrum gracile*	FWS	Nootka Sound
s31	*Chattonella* sp., *Dictyocha* spp., *Heterosigma akashiwo, Prorocentrum gracile*	FWS	Nootka Sound
s32	*Pseudopedinella* sp.?	FWS	Quatsino Sound
s33	*Chattonella* sp., *Pseudochattonella* sp., *Dictyocha* spp., *Cochlodinium fulvescens*	FWS	Quatsino Sound
s34	*Chattonella* sp., *Pseudochattonella* sp., *Dictyocha* spp., *Cochlodinium fulvescens*	FWS	Quatsino Sound
s35	*Heterosigma akashiwo*	FWS	Quatsino Sound

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

s36	*Heterosigma akashiwo*	FWS	Quatsino Sound
s37	*Heterosigma akashiwo,* nanoflagellates, big Raphidophyte	MC	Baynes Sound
s38	*Heterosigma akashiwo* (various morphotypes), nanoflagellates	FWS	Quatsino Sound
s39	*Chaetoceros* spp., *Skeletonema costatum, Dictyocha* spp., *Chrysochromulina* sp.?	FWS	Quatsino Sound
s40	Diatoms, *Dictyocha* spp., *Chrysochromulina* sp.?	FWS	Quatsino Sound
s41	Green algae, Cyanobacteria	FWS	Quatsino Sound
s42	*Heterosigma akashiwo*	FWS	Tofino Inlet
s43	*Heterosigma akashiwo*	FWS	Tofino Inlet
s44	*Chrysochromulina* sp.?	FWS	Kunechin Point
s45	*Chrysochromulina* sp.?	FWS	Quatsino Sound
s46	*Heterosigma akashiwo,* nanoflagellates, *Chattonella* sp.?, diatoms	MC	Quatsino Sound
s47	*Heterosigma akashiwo,* nanoflagellates, *Chattonella* sp.?, diatoms	MC	Quatsino Sound
s48	*Heterosigma akashiwo,* nanoflagellates, *Chattonella* sp.?, diatoms	MC	Quatsino Sound
s49	*Heterosigma akashiwo,* nanoflagellates, *Chattonella* sp.?, diatoms	MC	Quatsino Sound
s50	*Pseudochattonella* cf. *verruculosa*	PWS	Quatsino Sound
s51	*Pseudochattonella* cf. *verruculosa*	PWS	Quatsino Sound
s52	*Pseudochattonella* cf. *verruculosa*	PWS	Quatsino Sound
s53	*Pseudochattonella* cf. *verruculosa*	PWS	Quatsino Sound
s54	*Pseudochattonella* cf. *verruculosa*	PWS	Quatsino Sound

999

1000 ? = suspected species; PC = purchased culture; IC = isolated culture; MC = mixed culture; FWS
1001 = filtered water sample; PWS – pelleted water sample.

1002 **Table 2.** Overview of the algal taxa and qPCR primers included in this study.

Family	Genus	Species	Assay	Size (bp)	Citation for the primer	Forward Name	Forward Sequence	Reverse Name	Reverse Sequence	Probe Name	Probe Sequence
Bacillariophyceae	*Chaetoceros*	*tenuissimus*	C.ten	64	Toyoda et al., 2010	2–10_EF	TTG TGG AGA GGT ACG CTT GTC TT	2–10_ER	CCC TCA TAG GCA CCC TGT TC	2–10_Probe	CCT TAG CTT AAA TCT CT
	Pseudo-nitzschia	*multiseries*	P.mul	106	Delaney et al., 2011	Pmulti_F	AGATTTAACTGATGAACAAA	Pmulti_R	GTAACTATTACGTGGGTGT	Pmulti_P	CTATTAGCCGCGGTTTA
Cyanobacteria	*Cyanobacteria*	General	Cyan		Rinta-Kanto et al., 2005	CYAN-108-F	ACGGGTGAGTAACRCGTRA	CYAN-377-R	CCATGGCGGAAAATTCCCC	CYAN-328-P	CTCAGTCCCAGTGTGGCTGNTC
	Microcystis	spp.	Micr		Rinta-Kanto et al., 2005	Micr-184-F	GCCGCRAGGTGAAAMCTAA	Micr-431-R	AATCCAAARACCTTCCTCCC	Micr-228-P	AAGAGCTTGCGTCTGATTAGCTAGT
Dictyochophyceae	*Pseudochattonella*	*verruculosa*	P.ver1	227	Bowers et al., 2006	ChattaqFor	CCGTAGTAATTCTAGAGCTAATACRTGCA	ChattaqRev	AATTCTCCGTTACCCGTTAAAGCCAT	CVerrProbe	AATGGCGCGCAAGCGTGTATTATG
			P.ver2	239	Handy et al., 2006	Cv1561F	ATGCATACAGCGAGTCTAGA	Cv1780R	TCACTCCGAAAAGTGTCAAC	CvProbe	CAAGAGTACCCAGGCCTCTCGACC
Dinoflagellatephyceae	*Alexandrium*	*catenella*	A.cat	160	Hosoi-Tanabe and Sako, 2005	catF	CCTCAGTGAGATTGTAGTGC	catR	GTGCAAAGGTAATCAAATGTCC	Taq man cat	ATGGGTTTTGGCTGCAAGTGCA
		minutum	A.min	89	Toebe et al., 2013	Amin-F	ACATGGATAACTGTGGTAATTCTATAGCT	Amin-R	GTTGGTTCTGTAACTAATGACCACA	Amin-P	CATCCAAACCTGACTTC
		tamarense	A.tam1	131	Toebe et al., 2013	At-NA-1-F	CCACAGCCCAAAGCTCTTG	At-NA-1-R	CCATGAGGGAAATATGAAAAGG	At-NA-1-P	AACACTCCCACCAAGCA
			A.tam2	108	Toebe et al., 2013	At-ME-2-F	GCATGCCAAGAAAGAATCATTGT	At-ME-2-R	TGTATTTGCTGAACAAAAAGTAAACAGA	At-ME-2-P	AGCACAACAACCTCACCA
			A.tam3	82	Toebe et al., 2013	At-WE-3-F	TTGTGTGTGTCAGGGCTTGTA	At-WE-3-R	GAATGTGTCTGGTGTATCTGTT	At-WE-3-P	CAACCTCAAACACATGGA
			A.tam4	230	Hosoi-Tanabe and Sako, 2005	tamF	TGCTTGGTGGGAGTGTTGCA	tamR	TAAGTCCAAGGAAGGAAGCATC	Taqman tam	AGAGCTTTGGGCTGTGGGTGTA
		tamutum	A.tamu	90	Toebe et al., 2013	Amut-F	GAACATTTTGCAGCAACATTGTG	Amut-R	GTGGTACTTGTGTGTGTTCCTTTG	Amut-P	CTCAAGGAAGCAAACC
	Azadinium	*Obesum*	A.obe	74	Toebe et al., 2012	Aob134F	AGG GAT CGA TAC ACA AAT GAG TAC TG	Aob208R	AAA CTC CAG GGA CAT GGT AGT CTT A	Aob163	AAG ACA TTC GAC CTA CCG T
		poporum	A.pop	68	Toebe et al., 2012	Apop62F	GAT GCT CAA GGT GCC TAG AAA GTC	Apop148R	CCT GCG TGT CTG GTT GCA	Apop112	TTC CAG ACG ACT CAA A
		spinosum	A.spi	72	Toebe et al., 2012	Asp48F	TCG TCT TTG TGT CAG GGA GAT G	Asp120R	GGA AAC TCC TGA AGG GCT TGT	Aspin77T	CGC CCA AAA GGA CTC CT
	Cochlodinium	*polykrikoides*	C.pol	148	Park et al., 2009	CPITSF	CGGCAACCTTTGTCAAACA	CPITSR2	GGTTTGCTGATCTAACTTCATGTCT	CPITSP	CAACCGTGATACCCGCTAGCTTTGC
			Cpoly		Kamikawa et al., 2006	Cpoly-F	CCACACGGTGAGATTGTTGG	Cpoly-R	GCCACCGAAGTCGTTCG	Cpoly-P	CCACCAATCACCAATGACC
	Gyrodinium	*impudicum*	G.imp	105	Park et al.,	GIITSF	GAGTGCTTGTGCCT	GIITSR	CGCTTAGCCAGTT	GIITSP	CCTCAGGAGCGCCTCGA

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

	m				2009		CAGGAT		GGTTGAC		GAACA
		instriatum	G.ins	108	Handy et al., 2008	Gi 166F	GCACAAATTCCCA ACTTCGCGG	Gi 274R	GCTCGAATGATTC ATCGCCAGCA	GiProbe	CCAACACAGGCTCTGCC TGTTCTC
	Heterocapsa	*circularisquama*	H.cir		Kamikawa et al., 2006	HcircF	GTTTGCCTATGGGT GAGC	HcircR	CATTGTGTCAGGG AGGAG	Taq-Circ	CACCACAAGGTCATGAG GACACA
	Karenia	*Brevis*	K.bre	91	Gray et al., 2003		TGAAACGTTATTGG GTCTGT		AGGTACACACTTT CGTAAACTA		TTAACCTTAGTCTCGGGT A
		mikimotoi	K.mik1		Kamikawa et al., 2006	KmikiF	GTCTGGTAGCACTG CTTCA	KmikiR	GTCGGTTGCTGGT GCAAAA	Taq-miki	TGCTCCTGGCACCAACA AC
			K.mik2		Yuan et al., 2012	KMF	CTTTGTGTGTAACC TGTTGCTTTGT	KMR	TCAGCGGGTTTGC TTACCT	KMP	ACCTGTCCTCCTGTCTGC CACTTCATTTGT
	Karlodinium	*micrum/veneficum*	K.ven1	~150	Handy et al., 2008	Dino 1662F	CCGATTGAGTGWT CCGGTGAATAA	Euk B	GATCCWTCTGCAG GTTCACCTAC	KvProbe	CCGCTGCAACGTTCAGG AACTAAACACTG
		veneficum	K.ven2	120	Park et al., 2009	KVITSF 3	CTGTGAACTTCTTT GTGAGCTCTT	KVITSR 3	TAGCGATAGCTTC GCAGACA	KVITSP 3	AGGTGAATCCCAATGCT GCTCCACTA
	Pfiesteria	*piscicida*	P. pi	213	Bowers et al., 2000	107	CAGTTAGATTGTCT TTGGTGGTCAA	320	TGATAGGTCAGAA AGTGATATGGTA	Probe	CATGCACCAAAGCCCGA CTTCTCG
	Prorocentrum	*minimum*	P.min	325	Handy et al., 2008	Pm 200F	TGTGTTTATTAGTT ACAGAACCAGC	Pm 525R	AATTCTACTCATTC CAATTACAAGACA AT	PmProbe	CCGCCTGGTCCTTTGGTG ATTCATAATAAC
	Pseudopfiesteria	*shumwayae*	P.shu	78	Bowers et al., 2000	PSITSF1	TTGACGCATTGAG GCTATGG	PSITSR	AAACGGATGCAAA CGAGTTG	PSITSP1	CGGCTTCTCTGGCGACG AATCAC
Haptophyceae/ Prymnesiophyceae	*Prymnesium*	*Parvum*	P.par1	195	Manning and La Claire, 2010	FucoF	CTGCAGATGGTCTT CCTCATC	FucoR	AACCCTGGAGAGA CGAAGTTC	0	AAGCCTGACCTTGACGC TGATGAG
			P.par2	132	Zamor et al., 2012	PrymF	TGTCTGCCGTGGAC TTAGTGCT	PrymR-3	ATGGCACAACGAC TTGGT	PrymP	ACGTGTGCCGACGTGCT AGTAG
Raphidophyceae	*Chattonella*	*marina/ovata/antiqua*	C.mar1		Kamikawa et al., 2006	ChattF	GCTCTTTGTTCTCT GCATCC	ChattR	CAGTCCAATCACT GCAACC	Taq-Chatt	GCAACGGCAAGATGGAA CC
			C.mar2	224	Bowers et al., 2006	Cmarina For	GGTAGTTGCCGTAC ATTTTGCTCTT	Cmarina Rev	AAAAGTGGATTCA GCCGAAGCTTC	Cmarina Probe	TTGAGTTCAACGGGCGT GGTAGC
		subsalsa	C.sub1	156	Bowers et al., 2006	Subsalsa For	TTGGATTCCGACGG GC	Subsalsa Rev	ATATGCTTAAATT CAGCGGGTTTT	Subsalsa Probe	TTCGGCCAAGCACACAT TTCGGCCAAGCACACAT CCTC
			C.sub2	~350	Coyne et al., 2005	Cs 1350F	CTAAATAGTGTGG GTAATGCTTAC	Cs 1705R	GGCAAGTCACAAT AAAGTTCCAA	Raph Probe	CAACGAGTACTTTCCTTG GCCGGAA
	Fibrocapsa	*japonica*	F.jap	167	Bowers et al., 2006	Fjap490 For	TGAAAACGGCCCG TACACA	Fjap657Rev	CGGGAACAGCTCA TGATGT	Fjap578Probe	CGGCTGGACACGCTTCT G
	Heterosigma	*akashiwo*	H.aka1	228	Bowers et al., 2006	Haka127 For	AAAGGTGCGTGCT CAGTCGTGGT	Haka355 Rev	CAAAAGTCTTTTC ATCTTTCCCT	Haka222 Probe	TACGAGCCGTTTCCGAC GA
			H.aka2	~350	Coyne et al., 2005	Hs 1350F	CTAAATAGTGTCG GTAATGCTTCT	Hs 1705R	GGCAAGTCACAAT AAAGTTCCAT	Hs Probe	CAACGAGTAACGACCTT TGCCGGAA
			H.aka3		Kamikawa et al., 2006	HakaF	CTGCTGAAGGAAG CGATTG	HakaR	GGAACGCGCAGTC AAAG	Taq-aka	GGAGCGTTTCAACATGC GTTC

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 3. TaqMan assay amplification results (Ct values).

Sample #	Microscopy observations	A.tam1	Cyan	G.ins_	H.aka1	H.aka3	K.ven1	K.ven2	P.ver1
s01	*Alexandrium* cf *catenella*	8.0			24.2				
s02	*Chattonella subsalsa*								
s03	*Alexandrium* sp.	9.8			23.2				
s04	*Protoceratium reticulatum*	21.7	26.5		25.3				
s05	*Pseudo-nitzschia* spp.	25.1			26.1				
s06	*Heterosigma akashiwo*				25.2				
s07	*Heterosigma akashiwo*				28.6				
s08	*Heterosigma akashiwo*	36.2			18.6				
s09	*Heterosigma akashiwo,* big Raphidophyte (*Chattonella* sp.?), nanoflagellates				11.7				
s10	*Alexandrium* sp., flagellates, green algae	22.7	18.4		28.7				
s11	*Alexandrium* sp.	17.2			25.4				
s12	*Alexandrium tamarense*	8.7							
s13	*Alexandrium* sp.	8.4	24.4						
s14	*Pseudochattonella* cf *verruculosa*	25.7			26.6				14.0
s15	*Chaetoceros eibenii*	23.9			23.6	27.1			
s16	*Heterosigma akashiwo*				26.1				
s17	*Heterosigma akashiwo, Chattonella* sp.?								
s18	*Heterosigma akashiwo*				26.4				
s19	*Heterosigma akashiwo*				25.8				
s20	*Heterosigma akashiwo*				26.2				
s21	*Heterosigma akashiwo*				28.9				
s22	*Heterosigma akashiwo*				17.1				
s23	*Heterosigma akashiwo*				10.3				
s24	*Pseudopedinella* sp.	27.1							
s28	*Pseudopedinella* sp.						16.3		
s31	*Chattonella* sp., *Dictyocha* spp., *Heterosigma akashiwo, Prorocentrum gracile*		29.1	30.9	24.2		14.2	29.1	
s32	*Pseudopedinella* sp.?	35.2			17.6		22.9		
s33	*Chattonella* sp., *Pseudochattonella* sp., *Dictyocha* spp., *Cochlodinium fulvescens*	25.7	18.8		23.9		21.1		

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

s34	*Chattonella* sp., *Pseudochattonella* sp., *Dictyocha* spp., *Cochlodinium fulvescens*	21.2	16.9		20.5		13.4	29.2
s35	*Heterosigma akashiwo*				23.6			
s36	*Heterosigma akashiwo*				19.8			
s37	*Heterosigma akashiwo,* nanoflagellates, big Raphidophyte				13.7		25.6	
s38	*Heterosigma akashiwo* (various morphotypes), nanoflagellates				12.7			
s39	*Chaetoceros* spp., *Skeletonema costatum, Dictyocha* spp., *Chrysochromulina* sp.?							
s40	Diatoms, *Dictyocha* spp., *Chrysochromulina* sp.?	24.2		34.7	27.4		20.2	
s41	Green algae, Cyanobacteria						19.7	
s42	*Heterosigma akashiwo*	24.6	25.6					
s43	*Heterosigma akashiwo*	28.6	21.6					
s44	*Chrysochromulina* sp.?	24.4		31.7	27.2			
s45	*Chrysochromulina* sp.?	34.1			29.5		22.2	
s46	*Heterosigma akashiwo,* nanoflagellates, *Chattonella* sp.?, diatoms				22.1			
s47	*Heterosigma akashiwo,* nanoflagellates, *Chattonella* sp.?, diatoms				24.4			
s48	*Heterosigma akashiwo,* nanoflagellates, *Chattonella* sp.?, diatoms				25.6			
s49	*Heterosigma akashiwo,* nanoflagellates, *Chattonella* sp.?, diatoms				27.0			
s50	*Pseudochattonella* cf *verruculosa*	36.0					22.5	
s53	*Pseudochattonella* cf *verruculosa*		23.1		25.7			

Note: Full name for the primer abbreviation and its citation are listed in Table 2. Note that Ct's on the microfluidics BioMark platform are approximately 10 Ct values lower than on traditional single-assay platforms (Miller et al., 2016).

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1007 **Table 4.** Taxon levels and number of reads detected by the four different amplicons: 16S, 18S-
1008 Diatom, 18S-Dinoflagellate, and large subunit (LSU).
1009

Taxon level	16S		18S Diatom		18S Dino		LSU	
	Count	Reads	Count	Reads	Count	Reads	Count	Reads
Species	21	1281094	18	1947366	49	667599	60	1819902
Genus	27	229337	32	570709	52	1005795	46	103234
Family	11	41069	10	1095	23	104880	10	14705
Order	2	100	6	17002	18	96580	6	3295
Higher taxonomic rank	5	976679	8	29705	20	2222	9	6296
N/A, unknown	3	87534	4	122585	5	270013	5	291360
Total individual taxa	**69**	**2615813**	**78**	**2688462**	**167**	**2147089**	**136**	**2238792**

1010

1011 **Table 5.** Number of reads per taxonomic category and percent of total microalgae reads.

Row Labels	16S	18S Diatom	18S Dinoflagellates	LSU
Diatoms	30	291561	128820	49063
Dinoflagellate		557	1087060	630785
Raphidophytes	1229547	1993049	509885	1249934
Silicoflagellates		12272	5992	
Cryptophytes			366	594
Haptophytes		134	962	982
Microalgae (other)	95	238460	116431	4944
Cyanobacteria	101	14	21	
Bacteria	1076778			450
Fungi		30	606	3620
N/A, unknown	87534	122585	270013	291360
Other	221728	29800	26933	7060
Percent of total reads in microalgae taxa (%)	**47**	**94**	**87**	**86**

1012

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1013 **Table 6.** Number of reads for all detected harmful and potentially harmful algae species, genera,
1014 and families.

Taxonomic group	**16S**	**18S Diatom**	**18S Dino**	**LSU**	Amplicon returning maximum read
Alexandrium andersonii			97		18S Diatom
Alexandrium fundyense				549437	LSU
Alexandrium spp.		509	920011	14871	18S Dino
Alexandrium tamarense			4300	292	18S Dino
Total *Alexandrium* spp.	**0**	**509**	**924408**	**564600**	**18S Dino**
Chattonella spp.		152203	40	64380	18S Diatom
Chattonella subsalsa			58	28	18S Dino
Total *Chattonella* spp.	**0**	**152203**	**98**	**64408**	**18S Diatom**
Dinophysis parvula				13	LSU
Dinophysis spp.			119	111	18S Dino
Phalacroma rapa				44	LSU
Total *Dinophysis* and *Phalacroma* spp.	**0**	**0**	**119**	**168**	**LSU**
Karlodinium spp.			21	50	LSU
Karlodinium veneficum				27	LSU
Total *Karlodinium* spp.	**0**	**0**	**21**	**77**	**LSU**
Prymnesium kappa				20	LSU
Prymnesium spp.			10		18S Dino
Total *Prymnesium* spp.	**0**	**0**	**10**	**20**	**LSU**
Pseudo-nitzschia sp. B HAB-2017				33	LSU
Pseudo-nitzschia spp.		72978		11254	18S Diatom
Total *Pseudo-nitzschia* spp.	**0**	**72978**	**0**	**11287**	**18S Diatom**
Pseudochattonella farcimen		167			18S Diatom
Pseudochattonella spp.		1944	2104		18S Dino
Total *Pseudochattonella* spp.	**0**	**2111**	**2104**	**0**	**18S Dino**
Chattonellaceae			38486		18S Dino
Chrysochromulina spp.				220	LSU
Cochlodinium spp.				1016	LSU
Heterosigma akashiwo	1229451	1840799	471269	1185484	18S Diatom
Protoceratium reticulatum			145010	43109	18S Dino
Prymnesiaceae			11		18S Dino

1015

bioRxiv preprint doi: https://doi.org/10.1101/816926; this version posted October 24, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1016 **Table 7.** Evaluation of harmful taxa detection methods based on the present study: ** = best
1017 current technique(s) for taxa detection (at least one designated per taxon), * = technique that can
1018 detect taxa, 0 = no taxa detection, n/a = not applied.

Harmful taxa	qPCR	16S	18S-Diatom	18S-Dinoflagellate	LSU
Alexandrium tamarense complex	*	0	*	**	**
Chattonella spp.	0	0	**	*	*
Chrysochromulina spp.	n/a	0	0	**	*
Cochlodinium spp.	n/a	0	0	0	**
Cyanobacteria	*	**	0	0	0
Dictyocha spp.	n/a	0	0	0	0
Dinophysis and *Phalacroma* spp.	n/a	0	0	**	**
Heterosigma akashiwo	*	*	**	*	**
Karlodinium spp.	*	0	0	*	**
Protoceratium reticulatum	n/a	0	0	**	*
Prymnesium spp.	n/a	0	0	*	**
Pseudochattonella spp.	*	0	**	*	0
Pseudo-nitzschia spp.	n/a	0	**	0	*
Pseudopedinella spp.	n/a	0	0	0	0

1019