Metabarcoding reveals high genetic diversity of harmful algae in the coastal waters of Texas, Gulf of Mexico

Chetan. C. Gaonkar[1] and Lisa Campbell[1*]

[1]*Department of Oceanography, Texas A&M University, College Station, Texas 77843, USA*

*Corresponding Author: Lisa Campbell

email: lisacampbell@tamu.edu

Mailing Address:

Lisa Campbell

TAMU 3146 Eller O&M Building

College Station, TX 77843

Abstract

Environmental-DNA (eDNA) for metabarcoding is a rapid and effective means to investigate microplankton community composition and species diversity. The objective of this study was to examine the genetic diversity of the phytoplankton community in the Gulf of Mexico, with particular emphasis on harmful algal bloom species. Samples were collected at stations along the coast of Texas in September-October 2017 that were inundated by low salinity waters in the aftermath of Hurricane Harvey. Metabarcodes were generated from the eDNA targeting the V4 and V8-V9 regions of the 18S rDNA gene. Evaluation of the metabarcodes revealed an unexpectedly high number of harmful algal species during this short period, including five

that had not been documented in this region previously. A total of 36 harmful algal species could be differentiated based on V4 and V8-V9 metabarcode markers. Using a phylogenetic approach, the taxonomic resolution of each marker differed and not all species could be differentiated using solely one marker. The V4 region resolved species within some genera (e.g., *Heterocapsa*), while the V8-V9 marker was necessary to resolve species within other genera (e.g., *Chattonella*). In other cases, species differentiation within a genus required a combination of both markers (e.g., *Prorocentrum, Karenia*), or another marker will be needed to resolve all species (e.g., *Alexandrium*, *Dinophysis*). We conclude that no single marker can delineate all species, so it is recommended HAB monitoring programs use more than one marker. Overall, the observed diversity of HAB species along the Texas coast using metabarcoding exceeded reports from other parts of the world.

Keywords

genetic diversity, HAB, HTS-metabarcoding, phylogeny

Abbreviations

GoM, Gulf of Mexico, HAB, harmful algal blooms, HTS, high-throughput sequence, OTUs, operational taxonomic units, IOC-UNESCO, Intergovernmental Oceanographic Commission of UNESCO

1. Introduction

Marine phytoplankton include a diverse group of species that play a vital role in the ocean food web as primary producers. As such, the community composition and structure of the phytoplankton are an important factor in determining the functioning of marine food webs. Over

45 the last two decades, interest has focused on the diversity of protists and their role in marine
46 global biogeochemical cycles. Among the estimated >70,000 species of algae (Guiry, 2012),
47 only 153 are listed on the IOC-UNESCO Taxonomic Reference List of Harmful MicroAlgae
48 website, where HAB species are defined as "species producing toxins or toxic effects"
49 (Lundholm et al., 2009). This list includes species from six eukaryotic groups but does not
50 include species that cause harm due to biomass accumulation, mucus production or morphology
51 (e.g., setae). A recent review of the diversity HAB species on the coasts of the US found that
52 more than half of the 153 HAB species (98 species) are reported in US waters and most of these
53 (84 species) are found in the Gulf of Mexico (GoM) (Anderson et al., 2021). Historically,
54 *Karenia brevis* has been the primary HAB species of concern in the GoM (Steidinger, 2009) and
55 only in the past decade has *Dinophysis* emerged as a problem (Campbell et al., 2010). Given the
56 apparent increase in HAB events reported globally (Hallegraeff et al., 2021a), there are
57 increasing safety concerns for protecting human health as well as aquaculture and increasing
58 needs for improved methods of detection.

59 Most HAB monitoring programs are conducted using morphology-based methods, which are
60 labor-intensive and require taxonomic expertise. Some HAB species can be overlooked,
61 however, because of their small cell size or low abundance. In addition, results can sometimes be
62 misleading owing to the lack of distinct morphological traits to differentiate HAB vs non-HAB
63 species within the same genus, i.e., cryptic species. The inclusion of cryptic diversity and
64 presence of very small cells is known to impose challenges in accurate identification and
65 diversity estimates (Lundholm et al., 2006; Amato & Montresor, 2008; John et al., 2014).

66 The development of molecular-based methods and high-throughput sequencing (HTS) has
67 provided powerful tools for characterizing the diversity of the phytoplankton. In particular, the

metabarcoding approach, which incorporates molecular marker amplification and HTS, has been successfully employed to assess the species composition, diversity, and distribution in natural phytoplankton communities (Nagai et al., 2016; de Luca et al., 2019) and has facilitated the characterization and differentiation of cryptic species (Lundholm et al., 2006; Gaonkar et al., 2017; de Luca et al., 2021).

With the increasing demand for environmental monitoring programs, the trend has been a shift away from the classical morphological approach and toward DNA-based metabarcoding (Abad et al., 2016; Deagle et al., 2018; Pawlowski et al., 2018; Caracciolo et al., 2022). The most common marker genes used in phytoplankton community diversity studies are the V4 and V8-V9 hypervariable regions of the 18S rDNA gene. The preference for which marker to use for phytoplankton diversity studies has been debatable as both have advantages and disadvantages. Prior studies have shown that although both markers provided similar patterns in alpha diversity, V9 revealed a higher number of OTUs than V4 (Nanjappa et al., 2014; Tragin et al., 2018; Choi and Park, 2020). Contradictory results were observed for diatoms, however, where species diversity was higher using the V4 barcode compared to V8-V9. For example, in *Chaetoceros*, a collapse in the V9 terminal clade leads to a decrease in diversity, and the corruption of V9 forward primer due to introns (Gaonkar et al., 2018). Since both these markers have been used in monitoring of harmful algal diversity (Xu et al., 2017; Liu et al., 2020; Huang et al., 2021; Funaki et al., 2022), the objective here was to highlight the importance of choosing a marker based on the question addressed. Can a single marker based metabarcoding can differentiate and reveal the local HAB species diversity?

The present study was conducted during a 7-week period during Sept-Oct 2017 to assess HAB species diversity in the GoM on the Texas coast following Hurricane Harvey. Hurricane Harvey

was the wettest storm in recent history, resulting in 1.5m of rain in the Houston, TX area and producing an estimated 17 ± 5 km^3 of freshwater that was released from Galveston Bay onto the coastal shelf (Thyng et al., 2020). Coastal eutrophication has been linked with harmful algal bloom proliferation (Howarth 2008). Given the known impacts of terrestrial runoff on coastal waters following cyclones (Anglès et al., 2015; Fiorendino et al., 2021), the objective of this study was to assess the metabarcoding approach for estimating HAB species diversity and relative abundance. Because a single marker has not been identified for barcoding the phytoplankton, it was important to compare the resolution of the V4 vs. V8-V9 markers for HAB species detection and to characterize the potential HAB species using a phylogenetic approach. Moreover, results were used to compare HAB species richness in the GoM with time series in other regions of the world. Results have uncovered the genetic diversity of the HAB community in the GoM and will also provide relevant information for developing future monitoring programs. Given the perceived global increase in HAB events (Hallegraeff et al., 2021b), improved identification of HAB species is needed.

2. Materials and Methods

2.1. Sample site and collection

This study was conducted in the Gulf of Mexico along the Texas coast after the passage of Hurricane Harvey on August 25, 2017. A total of 36 samples were collected on 5 dates during September to October 2017 (Supplemental Fig. 1). Surface seawater (~3 to 10 m) was filtered on 5.0 μm 47 mm diameter Durapore filters (Millipore, USA) in triplicate for DNA extraction using the AllPrep DNA/RNA MiniKit (Qiagen, USA) following the manufacturer's instructions.

The Texas Observatory for Algal Succession Time series (TOAST) is a network of Imaging FlowCytobots (IFCB) on the Texas coast. The IFCB at Port Aransas has been deployed since 2007 (Campbell et al., 2010; Fiorendino et al., 2021). Images from TOAST were used to identify HAB species for this study (*https://toast.tamu.edu/timeline?dataset=PortAransas*).

2.2. Metabarcoding and HTS sequencing

DNA concentration and quality were evaluated using a Nanodrop spectrophotometer (Thermo Fisher Scientific Inc, USA) and normalized to 5 ng/µl concentration for the amplicon library construction. The hypervariable regions (V4 and V8-V9) of the 18S rDNA gene were amplified using customized primers: V4f CCAGCASCYGCGGTAATTCC; V4r ACTTTCGTTCTTGAT V8-V9f ATAACAGGTCTGTGATGCCCT; V8-V9r CCTTCYGCAGGTTCACCTAC (Bradley et al., 2016; Kozich et al., 2013) following the modifications of Gaonkar et al. (2020). Amplifications were performed in triplicate 25 µl reactions for each sample consisting of ca. 5 ng of extracted DNA, 0.5 µM primers and 1x GoTaq® Green Master Mix (Promega, USA). PCR cycles were set at initial denaturation of 95 ºC for 300 s, followed by 10 cycles of denaturation at 95 ºC for 30 s, annealing at 47 ºC for 45 s, extension at 72 ºC for 60 s, and subsequently 15 cycles of denaturation at 95 ºC for 35 s, annealing at 57 ºC for 40 s, extension at 72 ºC for 60 s, with a final extension step of 10 min at 72 ºC for both markers. A negative control was included containing only Optima water (Fisher Scientific, USA).

Replicate PCR products for each sample were merged and purified using a DNA Purification SPRI Magnetic Beads (ABM, Canada) and quantified using NanoDrop 3300 (Thermo Fisher Scientific). Next, all samples were pooled to an equimolar concentration and repurified. This product was analyzed using an Agilent Fragment Analyzer Systems DNA Analysis Kits - NGS

Fragment Kit (1-6000bp) and sequenced on an Illumina MiSeq Platform using a v3 kit (2x300bp). The amplicon library sequencing was performed by the Genomics and Bioinformatics Services at Texas A&M University (*https://www.txgen.tamu.edu*) using custom designed primers (Bradley et al., 2016).

2.3. HTS-data processing and analysis

MiSeq paired end reads of 2x300 bp were analyzed using mothur v1.39.0 (Schloss et al., 2009) using the protocol outlined by Piredda et al. (2017). The resulting contigs with three or more reads were BLASTed (blastn) against PR2 database v4.12.0, for taxonomic annotation (Guillou et al., 2013; del Campo et al., 2018). Only annotations with molecular operational taxonomic units (OTUs) similarity of ≥ 90% and query coverage ≥ 70% against the HAB reference sequence (IOC-UNESCO list) were included in this study. Taxonomically assigned sequences (OTUs) annotated as HABs used in the study were submitted to Biological & Chemical Oceanography Data Management Office (BCO-DMO) under the project number 715170.

2.4. HAB delineation based on phylogenetic approach

OTUs recovered by clustering followed by BLAST may differ depending on the % similarity threshold level used. These potential HAB OTUs were compared against the reference sequences using a phylogenetic approach. Following the BLAST analysis and screening procedure, the putative OTUs were aligned with the IOC-UNESCO designated HAB reference sequences from GenBank using MAFFT v7.245 (Katoh and Toh, 2008) and manually corrected using SeaView v4.5.4 (Gouy et al., 2010). A maximum likelihood (ML) tree was generated using RAxML v2.0.5 (Stamatakis, 2014) using a GTRGAMMA substitution model with 1000 replicate for bootstrap analysis. Those OTUs that were recovered outside the clades were removed and the

ML tree was generated again. The resulting OTUs which were represented with a reference HAB sequence within or along the clades were considered validated HAB species.

3. Results

3.1. Composition and diversity of the microplankton community in Gulf of Mexico, Texas coast

From the BLAST results, after the initial screening criterion of ≥90% similarity and ≥70% sequence length coverage, a total of 16,918 (V4) and 10850 (V8-V9) OTUs were recovered that were assigned to microplankton (excluding Metazoans). To examine the diversity of the microplankton in the northern Gulf of Mexico along the Texas coast, the annotated OTUs were first clustered at subkingdom level: Harosa (92%), Hacrobia (4.6%), Viridiplantae (1.9%), Eozoa and others (1.5%). Among these, Harosa (Alveolata, Stramenopiles and Rhizaria) was the most abundant subkingdom largely comprised of Dinoflagellata (70.4%) and Ochrophyta (16.9%). Hacrobia which included cryptophytes, haptophytes, picozoa and others, was the second most abundant group (Fig. 1). Viridiplantae (Archaeplastida) and Eozoa (protozoan and amoebozoans) were minor components with only a few thousand reads in both V4 and V8-V9 markers. Based on the BLAST results, only 5 and 3.5 % of the OTUs were represented as potential HABs with V4 and V8-V9 markers, respectively (Fig. 1).

3.2 Diversity of HAB OTUs

Based on the BLAST analysis following the criteria of similarity (≥90% similarity and ≥70% sequence length coverage), we found 835 OTUs (n=163,689) with the V4 marker and 393 OTUs (n=125,266) with the V8-V9 marker annotated as HAB species. Among the 835 OTUs annotated as HAB species with the V4 marker, 50 OTUs were classified as diatoms, 744 as dinoflagellates,

179 26 as raphidophytes, and 15 as haptophytes. *Karenia* and *Karlodinium* shared the highest number
180 of OTUs and were most abundant among the dinoflagellates, while *Pseudo-nitzschia* was
181 dominant in diatoms in both the markers. Classifications to species level using ≥95% similarity
182 and ≥95% sequence length coverage to minimize the number of OTUs with lower similarity and
183 shorter sequence length (Gaonkar et al. 2020) revealed six species in diatoms, 19 in
184 dinoflagellates, three in haptophytes and three in raphidophytes, for a total of 31 species
185 (Supplemental Table 1).

186 Among the 394 OTUs identified with the V8-V9 marker, 59 OTUs were classified as diatoms,
187 307 as dinoflagellates, 17 as raphidophytes, and 11 as haptophytes. Again, clustering OTU
188 classifications to species level revealed eight species in diatoms, 16 dinoflagellates, three in
189 haptophytes and two in raphidophytes, for a total of 29 species. Exceptions include
190 *Prorocentrum lima*, which was only 95.6% (V4) and 93.8% (V8-V9) similar to the reference
191 sequence, and *Karenia brevis*, which was 91.3 % (V8-V9) similar to the reference sequence.
192 Only 23 HAB species were common among both the V4 and V8-V9 markers (details can be
193 found in Supplemental Table 2). Moreover, five HAB species that had not been documented in
194 this region previously were identified in this study (Table 1).

195 **3.3. Diversity of HAB species illustrated with a phylogenetic approach**

196 A phylogenetic approach was carried out to verify annotations because some of the OTUs had
197 low sequence similarities. The dominant OTUs for each potential HAB species were aligned
198 along with the reference sequences for HAB species listed by the IOC-UNESCO for the
199 phylogenetic analysis. The resulting ML-tree for the V4 marker revealed at least 26 HAB species
200 and for the V8-V9 marker the number was 24 potential HAB species (Fig. 2A&B). The putative
201 OTUs with similarities <99% were removed from the final trees. Based on the V4 and V8-V9

phylogenies for the HAB species, it was evident that the three major groups, dinoflagellates, diatoms and raphidophytes, were well resolved (Fig. 2A&B). The clade structure appeared to be basically the same in the two phylogenies. The ingroup dinoflagellates formed a clade with good bootstrap values (71 and 100 in V4 and V8-V9, respectively) but showed weak support for the basal topology in the V4 compared to the V8-V9 marker. The species belonging to the genera *Karenia*, *Karlodinium*, *Amphidoma*, *Heterocapsa,* and *Prorocentrum* were poorly resolved. Moreover, based on the two phylogenies it was evident that *Prorocentrum* species are paraphyletic.

3.4. Resolution of species based on marker region

The generated HAB reference dataset for identification and delineation of these species included 100 V4 and 98 V8-V9 reference sequences. A total of 422 positions were present in the V4 marker of which 281 formed a distinct alignment pattern, while there were 340 positions in the V8-V9 region of which only 223 formed a distinct alignment pattern. Among the *Karenia*, *Karlodinium*, *Azadinium*, *Amphidoma*, *Heterocapsa* and *Dinophysis* genera, only 29 (V4) and 52 (V8-V9) distinct alignment patterns were evident, which supports the low-resolution power and poor bootstrap values at the basal clade of the V4 marker. Most of the reference sequences in the V8-V9 region are not complete in the 3′-end, which explains the poorly supported ramification.

Based on the phylogenetic approach it was evident that the resolution power of the V4 and V8-V9 marker genes differed (Fig. 2A&B, Fig. 3A-J). Overall, the V4 marker had better resolution for the diatoms and prymnesiophytes: *Halamphora coffeaeformis, Nitzschia bizertenisis* and *Pseudo-nitzschia spp*. were distinguished, as were *Phaeocystis pouchetii* and *P. globosa* (Fig. 2A vs. 2B). Among the dinoflagellates, the V4 marker was able to distinguish the toxic *Heterocapsa circularisquama* from the non-toxic species (Fig. 2A&B). The ML-tree clearly shows species

225 within this genus can be resolved using the V4 marker while V8-V9 marker fails to differentiate
226 the non-toxic species vs the toxic species.

227 Overall, the V8-V9 marker region appeared to resolve the clade structure for the dinoflagellates
228 with more support. In the genus *Prorocentrum* only a few species are toxic; however, the V4
229 marker failed to delineate the toxic vs non-toxic species, while V8-V9 marker was successful in
230 differentiating *P. cordatum* (Fig. 3E&F). Similarly, only several species within the *Dinophysis*
231 genus could be delineated (*D. acuta, D. fortii,* and *D. infundibulus)* (Fig. 3B). The V8-V9
232 barcode was also able to distinguish *Karlodinium jejuense* and *K. veneficum* (Fig. 3H) and
233 *Chattonella antiqua* and *C. marina* (Fig. 3B).

234 In several cases to delineate species within a genus, a combination of both markers was
235 necessary because the resolution power of each marker individually was low. This was evident
236 for the genus *Alexandrium*. Although a number of species could be distinguished with either the
237 V4 or V8-V9 marker (e.g., *A. affine, A. andersonii, A. australiense, A. ostenfeldii, A. taylori*), the
238 V4 marker was required to differentiate *A. catenella* and *A. pacificum*; whereas a combination of
239 the two were needed to *resolve A. minutum, A. tamutum* and *A. insuetum* (Fig. 3C&D). Neither
240 marker was able to distinguish *A. hiranoi* and *A. pseudogonyaulax* (Fig. 3C&D). This difficulty
241 was also the case for *Karenia*. From the ML-tree generated with the V4 marker, *K. mikimotoi*
242 and *K. brevis* were identical, while *K. selliformis* formed a distinctive clade (Fig. 3G). In
243 contrast, from the V8-V9 ML-tree, the two species *K. mikimotoi* and *K. brevis* could be
244 differentiated, but the V8-V9 marker failed to discriminate *K. selliformis* from *K. mikimotoi* (Fig.
245 3H). Similarly, the HAB species *Azadinium poporum* and *A. spinosum* can be differentiated from
246 the non-HAB *Azadinium* species using a combination of V4 and V8-V9 markers (Fig. 2A&B).

In other cases, either the V4 or V8-V9 marker successfully differentiated the HAB species: *Gymnodinium catenella*, *Gonyaulax spinifera*, *Lingulodinium polyedra, Margalefidinium polykrikoides*, *Polykrikos hartmanii*, *Pyrodinium bahamense* (Fig. 2A&B). For the genera *Ostreopsis* and *Gambierdiscus,* either the V4 or the V8-V9 marker can be used to distinguish the 14 toxin-producers among the 16 species in the genus *Gambierdiscus* and the 7 toxin-producers among the 11 species of *Ostreopsis* (Fig. 3I&J).

3.5 Global distributions

Datasets from other coastal time series reporting HAB species using microscopy and metabarcoding were compiled to compare with this study in the Gulf of Mexico (Table 1). There are 84 known HAB species documented in the Gulf of Mexico region by microscopy, a total of 36 HAB species were observed on the Texas coast in just the short 7-week period of this metabarcording study (Table 1). Of these, 31 were previously known and an additional five species were identified in the Gulf of Mexico for the first time (*Alexandrium hiranoi*, *A. pacificum*, *Ampidoma languida*, *Nitzschia bizertensis*, and *Prymnesium polylepis*). In comparison with the other metabarcoding studies, the overall HAB diversity in this study was similar to Bohai Sea and the Changjiang estuary, but few species were common (Table 1). Filter cut-offs and sequencing depth differed from this study (Supplemental Table 3). Overall, 46 HAB species have been observed at TOAST from Imaging FlowCytobot images and HTS-metabarcoding between 2007 -2021 (Table 1). From images and microscopy, both the Gulf of Mexico and the Mediterranean Sea regions had the highest number of HAB species reported (84 species for both). In fact, the Mediterranean Sea had the most species in common with this study and the Gulf of Mexico overall.

4. Discussion

Among the tens of thousands of known phytoplankton species, only 153 species are known to produce toxins and are representative of HAB diversity (Lundholm et al., 2009). In recent years, metabarcoding has been proposed as an effective and reliable tool for measuring biodiversity of the phytoplankton, as well as HAB monitoring. This approach has shown promise for detecting hidden diversity (e.g., de Varges et al., 2015; Gaonkar et al., 2020), but the transition from classical light microscopy to metabarcoding for studies of HAB diversity is recent (Nagai et al., 2017; Xu et al., 2017; Liu et al., 2020; Huang et al., 2021). It can be seen from the results presented here that the choice of the marker to assess diversity may depend on the species of interest and the resolution power of the marker. Moreover, a phylogenetic approach should be included to confirm species identification, rather than relying on BLAST, to provide better resolution and accuracy.

4.1 Selection of appropriate HTS marker to delineate HAB species

One of the main advantages of employing HTS-metabarcoding is the ability to distinguish species, including the cryptic ones (e.g., Slapeta, 2006). Some species can be differentiated with a single marker, while others require a combination of both V4 and V8-V9 markers. For an HTS-based monitoring program, the choice of the marker to be used should be based on the target species of interest.

There are examples where the choice of the marker used for metabarcoding provides essential information. In case of *Heterocapsa*, characterizing the species within this genus morphologically is a challenging and difficult task. Based on the phylogenetic resolution, only the V4 marker can differentiate the species (Fig. 2A&B). This is important because not all members of the genus *Heterocapsa* are toxin producers; of the 19 species (WoRMS, 2022), only *H. circularisquama* is associated with mortality in bivalves (Horiguchi, 1995). The V8-V9

marker alone was useful in only a few cases, but it allowed the toxic *Karlodinium* species to be identified. Only 6 of the 12 *Karlodinium* species are reported to be toxic (Lundholm et al., 2009; WoRMS, 2022), but there are only reference sequences available for the toxic species *K. veneficum* and *K. australe*. Similarly, for *Dinophysis*, of the 123 documented species, only 10 are reported to be toxic (WoRMS,2022; Lundholm et al., 2009); however only *D. acuta* and *D. fortii* and *D. infundibulus* are potentially distinguishable from other species with the V8-V9 marker. Overall, most toxic species of *Dinophysis* cannot be distinguished with either marker (Fig. 3A&B) and a different target gene will be needed. In contrast, the two toxin-producing *Phalacroma* species (Lundholm et al., 2009) can be distinguished from the other 10 non-toxin producing species (WoRMS, 2022) with both V4 and V8-V9 markers (Fig. 3A&B).

In some cases, using both V4 and V8-V9 markers is necessary to identify target HAB species. For example, the genus *Karenia* includes 10 species, 9 of which are fish killers in coastal waters globally (WoRMS, 2022; Lundholm et al., 2009). In the GoM *Karenia brevis* is one of the most extensively studied HAB species due to its impact on ecosystem and human health (Brand et al., 2012), but there are six recognized species known to coexist. Metabarcoding results revealed that members of this genus are challenging to differentiate using a single marker. If the V4 marker was used, the presence of *K. mikimotoi* could lead to a false positive result if only *K. brevis* was present; the V8-V9 marker would be required for differentiation. This highlights the importance of marker choice.

4.2. Reference dataset and taxonomically validated sequences

Taxonomically validated reference barcodes play a key role in identification and delineation of the species diversity in a community. One important and necessary component of the metabarcoding approach is the availability of the taxonomically validated reference barcode

316 sequences. Lack of reference sequences will lead to misrepresentation of the species diversity
317 and decrease the count of the species present within a genus. Of the 153 HAB species listed by
318 the IOC, only 100 reference sequences were available for the V4 region and 98 for the V8-V9
319 region. This lack of reference sequences for ~one third of the species on the IOC list (Lundholm
320 et al., 2009) could lead to an underestimation of the potential HAB diversity. The lack of
321 reference sequences is also true for other protist groups (Santoferrara et al., 2020). If species are
322 not well represented in the reference dataset, the alpha and beta diversity estimates of a
323 community will be seriously underestimated.

324 Another purpose for creating a comprehensive reference dataset is to understand the genetic
325 variance among geographically distributed species. In the present study, only one representative
326 barcode for each species was used in the phylogenetic analyses, although peripheral
327 metabarcodes were also found. Peripheral metabarcodes differ by one or just a few bases from
328 the potential HAB reference sequence, which may be the result of geographic variation or
329 heterogeneity in the sequence. In this study, the dominant metabarcode was 100 % identical to
330 the reference sequence in ~half of the sequences. In the other cases (see 4.3 below), another less
331 abundant OTU was identical, which supports the conclusion for geographic variation or
332 heterogeneity in HAB species (Supplemental Table 1 and 2).

333 **4.3. Phylogenetic approach vs. BLAST analysis**

334 Many previous studies have relied on BLAST-based assumptions for taxonomic annotation to
335 explore species diversity (de Varges et al., 2015; Piredda et al., 2017; de Luca et al., 2019;
336 Gaonkar et al., 2020); however, there are limitations to this approach. One of the main
337 drawbacks of inferring species identifications only from BLAST results is the lack of visual

inspection to compare a sequence to the reference. Some reference sequences present in the PR2 database are curated sequences and some are from environmental clones (del Campo et al., 2018). For example, an OTU was annotated as *Pyrodinium bahamense* (AB936750) with 99.47% similarity and 100% length coverage based on BLAST with PR2 database. When checking the alignment visually, it was evident that the sequence was not *P. bahamense* and was not close to the clade including *P. bahamense*. To investigate this further, this sequence was blasted against GenBank and showed a 99.48% similarity, 100% length coverage, and 0.0 e-value with *Goniodoma polyedricum* (KM886380) and 93.35% similarity, 55% length coverage, and 0.0 e-value with *P. bahamense* (AB936750). Obviously, there can be errors in the reference databases. If only a BLAST approach had been used with the given reference database, results may have included errors.

At times, the most dominant OTU in a sample may not be identical to the reference HAB sequence, which may be due to geographic distance or sequence reading error. In this study, reference sequences from GoM were utilized; however, in cases of where GoM sequences were not available, references from other US coasts were chosen. If both were unavailable, a full length taxonomically validated 18S rDNA reference sequence from GenBank was used to generate phylogenies. The reason for prioritizing locally obtained reference sequences is that these are expected to match the dominant OTU, irrespective of species rarity (Gaonkar et al., 2020). An example for such a scenario in this data is an OTU that was identical to *Polykrikos hartmanii* (AY421789) but was not the dominant OTU in the list of OTUs matched with this reference sequence. The dominant OTU was one bp different, which could be a geographic variation because the reference sequence was generated from a Korean strain. Among the

360 *Karenia*, *Karlodinium*, *Amphidinium*, *Heterocapsa* and *Dinophysis* genera, the resolution power
361 of both markers is low, which can lead to confusion if results are not validated properly.

362 Phylogenetic analysis to delineate the toxin-producing HAB species, as defined by the IOC-
363 UNESCO, can markedly improve the OTUs recovered from the BLAST analysis for
364 metabarcoding. The advantage of using the phylogenetic approach is that it rectifies the
365 uncertainties inherent in the BLAST results. An example of this uncertainty is seen in the
366 *Karenia mikimotoi*. A BLAST search in GenBank revealed that the same OTU had multiple hits
367 with ≥ 99% similarity to *Karenia, Karlodinium*, *Azadinium*, and *Amphidinium*. When the
368 alignment was visually inspected for the sequence base differences, it was evident that the base
369 differences were present at different positions. This explains why the phylogenetic approach is
370 favored over the BLAST approach for species identification. The BLAST approach can still
371 provide preliminary information for HAB species identification. Examples include *Amphidoma*
372 *languida* and *Prorocentrum lima,* which were only identified by BLAST as their closest
373 representative reference HAB sequence (97.91 % similarity with *A. languida* reference sequence
374 and 95.56% similarity with *P. lima* reference sequence). Since these two OTUs had similarities
375 <99%, they did not meet the criterion of defining HAB species, so these OTUs were not included
376 in the phylogenies.

377

378 **4.4. Diversity of HAB species in the Gulf of Mexico, Texas coast via HTS-metabarcoding**

379 The Gulf of Mexico is a semi enclosed ocean basin that is one of the most productive US
380 aquatic ecosystems and supports major commercial and recreational fisheries. Global warming
381 and anthropogenic eutrophication have led to major impacts on coastal ecosystems. With climate

382 change, the expectation is that HAB species will increase their geographic range into new areas
383 that may facilitate HAB initiation and maintenance (see Wells et al., 2015; Gobler et al., 2020).
384 Additionally, with climate change there is potential for an increased frequency and intensity of
385 hurricanes (Emanuel, 2017). Increased freshwater discharge conditions after hurricane landfalls
386 appear to favor dinoflagellates in the GoM (Fiorendino et al., 2021). From this perspective, the
387 HAB diversity reported during this short period following Hurricane Harvey is revealing.

388 Most of the HAB diversity explored via metabarcoding in this study could not be easily
389 characterized based on traditional morphotaxonomic observation by light microscopy. First,
390 some HAB species share similar morphological traits with non-HAB species. For example, the
391 species in the genus *Heterocapsa* are all small cells with similar morphological traits, so are
392 difficult to distinguish (Iwataki, 2008), while only *H. circularisquma* is a toxic species. Other
393 examples include members of the genera *Azadinium* and *Phaeocystis*. Additionally, among the
394 *Pseudo-Nitzschia* species only ~half are toxic. Secondly, the morphological variation within
395 *Karenia* species can make distinguishing among species of this genus difficult. Third, the small
396 size and nondescript morphology of pico- and nano-plankton (< 5µm), such as *Karlodinium*
397 *veneficum*, make these cells challenging to identify by light microscopy. Therefore, these groups
398 of hidden species often are characterized only because of HTS-metabarcoding (Chen et al.,
399 2019). The small Pelagophyceae, such as *Aureoumbra lagunensis,* have been reported in the
400 GoM, but were not detected in this study. Since the extended brown tide bloom in the 1990s
401 (Buskey et al., 2001), the range of *A. lagunensis* has expanded to include other regions of Texas,
402 Florida, and Cuba (Hall et al., 2018); however, because the focus of the HTS-metabarcoding was
403 eukaryotes in the >5µm size fraction, these smaller sized species and HAB cyanobacterial
404 species were not represented in this study. Future studies should consider the target HAB species

405 when selecting filters for size fractionation. Alternatively, given the short duration of this study,
406 *A. lagunenis* may have been absent at this time, as was noted for several other HAB species
407 previously recorded at TOAST (Table 1).

408

409 **4.5. Global diversity of HABs**

410 The number of reported HAB events globally have increased in the last three decades, based on
411 the *Harmful Algal Event Database, HAEDAT* (*http://haedat.iode.org*). One explanation is the
412 intensification in monitoring effort (Hallegraeff et al., 2021a, b). But with global climate change
413 and continuing interference of anthropogenic activities, impacts on the productivity and
414 biodiversity of aquatic ecosystems are increasing (Heisler et al., 2008; Gobler, 2020).
415 Consequently, one of the intriguing questions currently addressed is whether there will be a
416 continued increase in HAB events. From HAB monitoring programs there appears to be growing
417 concern due to the negative effects of changes in environmental conditions on the diversity. Most
418 obvious is that many HAB species favor warmer waters (Gobler et al., 2017).

419 A comparison of the number of toxin-producing HAB species in this GoM study with coastal
420 time series data from other parts of the of the world found the highest species diversity to be
421 observed in the GoM, Mediterranean Sea and N. Atlantic (Table 1). The Mediterranean Sea and
422 the GoM are similar in terms of latitude, geographical characteristics, and environmental
423 parameters, so the large number of HAB species in both regions is not unexpected. The Long-
424 Term Ecological Research station in the Gulf of Naples has been assessing phytoplankton
425 diversity and distribution for close to four decades (Zingone et al., 2019). At this station, a total
426 of 54 HAB related species have been identified since 1984 and of this total, 22 species were

found to be in common with HAB species observed in this study (Cipolleta et al., 2021). Note, however, this compares ~40 years with 7 weeks of observations. Similarly, HAB diversity in the North Atlantic from 1989-2019 revealed presence of 74 toxic HAB species of which 23 species were common to the GoM (Bresnan et al., 2021). The diversity of HAB species in Australian and New Zealand waters included 43 toxic species, but only 13 species were common to the GoM (Hallegraeff et al., 2021c). In all the other long-term studies in the coastal seas of North Europe, the North Pacific, and the Gulf of Oman and Arabian Sea, observations were based on morphotaxonomic monitoring, and fewer than 20 toxin-producing species were found to be in common with the GoM (Al-Azari et al., 2015; Karlson et al., 2021; McKenzie et al., 2021; Sakamoto et al., 2021).

Most of the previous time series reports were based on monitoring data generated from morphological observations by light microscopy. Only recently has HTS-metabarcoding been investigated as an approach for monitoring. HTS-metabarcoding studies conducted in the Bohai Sea (Huang et al., 2021) and Changjiang estuary (Cui et al., 2021) have reported a similar number of HAB species as reported here for the GoM. Both of these prior studies were also short term (days to months); however only 16 species were common to GoM. In the Changjiang estuary, China, considered a region with frequent HAB occurrences, only 13 were common to the GoM.

HTS-metabarcoding has been compared with morphotaxonomy in a number of previous studies (e.g., Deagle et al., 2018); however, the lack of taxonomically validated reference sequences makes assessments difficult. In some cases light microscopy may still be preferred over V4 or V8-V9 markers (e.g., some species of *Karenia, Alexandrium, Dinophysis*); however, in other cases HTS-metabarcoding may be preferred (e.g., *Heterocapsa*, and benthic dinoflagellates,

Gambierdiscus, Ostreopsis). Still, there remain a number of genera for which more reference sequences are needed to provide more definitive species identifications (e.g., *Dinophysis, Phalacroma*, and *Pseudo-Nitzschia*).

5. Conclusions

Metabarcoding is a valuable tool to assess the species richness of the plankton. When applied for HAB species detection, it allows for a more detailed understanding of diversity than traditional light microscopy because it allows for the identification of rare and cryptic species. The choice of markers, however, is crucial. Results from HTS of the V4 and V8-V9 regions of 18S rDNA indicated no universal single marker is appropriate for all HAB species identification and in many cases a combination of two markers is recommended. Moreover, a phylogenetic approach to confirm species identification should be used instead of just relying on BLAST results. A key difference between the BLAST based similarity approach and phylogeny is that the latter is much more flexible for taxonomic resolution without the need for subjective decisions (Jamy et al., 2020). The clustering method classifies OTUs that include a few bp differences (which may-or-may not be in the exact same sequence position) to the same predetermined rank, whereas the phylogenetic approach allows distinctions based on sequence differences. Another benefit of the phylogenetic approach is that it can also allow discovery of new species that are lacking reference sequences. BLAST analysis is beneficial when taxonomic annotation is sought for a higher taxonomic level, but results need to be validated when describing at the species level. There is a need for taxonomically validated reference sequences for all HAB species, especially including geographic variants. From this study, HAB diversity in the Gulf of Mexico was found to be higher than expected for the limited sampling period (7 weeks) and, importantly, 5 species were documented that had not previously been reported in the GoM.

473 Acknowledgements

474 We thank the captain and crew of the *R/V Pt. Sur*, S.F. DiMarco, J.M. Fiorendino, and D.W.
475 Henrichs for their assistance in sample collection. This work was supported by a grant from the
476 National Science Foundation, OCE-1760620 to L.C. Any opinions, findings and conclusions or
477 recommendations expressed in this material are those of the authors and do not necessarily
478 reflect the views of the National Science Foundation. The authors declare that they have no
479 conflicts of interest.

References

Abad, D., Albaina, A., Aguirre, M., Laza-Martinez, A., Uriarte, I., Iriarte, A., Villate, F., Estonba, A., 2016. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Mar. Biol. 63(7), 1-13.

Al-Azari, A.R., Al-Hashmi, K.A., Al-Habsi, H., Al-Azri, N., Al-Khusaibi, S. 2015. Abundance of harmful algal blooms in the coastal waters of Oman: 2006–2011. Aquatic Ecosystem Health and Management 18(3), 269-281.

Anglès, S., Jordi, A., Campbell, L., 2015. Responses of the coastal phytoplankton community to tropical cyclones revealed by high-frequency imaging flow cytometry. Limnol. Oceanogr. 60(5), 1562-1576.

Amato, A., and Montresor, M., 2008. Morphology, phylogeny, and sexual cycle of *Pseudo-nitzschia mannii* sp. nov. (Bacillariophyceae): a pseudo-cryptic species within the *P. pseudodelicatissima* complex. Phycologia 47, 487–497.

Anderson, D.M., Fensin, E., Gobler, C.J., Hoeglund, A.E., Hubbard, K.A., Kulis, D.M., Landsberg, J.H., Lefebvre, K.A., Provoost, P., Richlen, M.L., Smith, J.L., Solow, A.R., Trainer, V.L., 2021. Marine harmful algal blooms (HABs) in the United States: history, current status and future trends. Harmful Algae 102, 101975.

Bradley, I.M., Pinto, A.J., Guest, J.S., 2016. Design and evaluation of Illumina MiSeq compatible, 18S rRNA gene-specific primers for improved characterization of mixed phototrophic communities. Appl. Environ. Microbiol. 82(19), 5878-5891.

Brand, L.E., Campbell, L., Bresnan, E., 2012. *Karenia*: The biology and ecology of a toxic genus. Harmful Algae 14, 156-178.

Bresnan, E., Arévalo, F., Belin, C., Branco, M. A., Cembella, A. D., Clarke, D., ... and Enevoldsen, H. (2021). Diversity and regional distribution of harmful algal events along the Atlantic margin of Europe. Harmful Algae 102, 101976.

Buskey, E.J., Liu, H.B., Collumb, C., Bersano, J.G.F., 2001. The decline and recovery of a persistent Texas brown tide algal bloom in the Laguna Madre (Texas, USA). Estuaries 24(3), 337-346.

Campbell, L.,Olson, R. J., Sosik, H. M., Abraham A., Henrichs, D. W., Hyatt, C.J., Buskey, E. J. 2010. First harmful *Dinophysis* (Dinophyceae, Dinophysiales) bloom in the US is revealed by automated imaging flow cytometry. J. Phycol. 46 (1), 66-75

Caracciolo, M., Rigaut-Jalabert, F., Romac, S., Mahe, F., Forsans, S., Gac, J.P., Arsenieff, L., Manno, M., Chaffron, S., Cariou, T., Hoebeke, M., Bozec, Y., Goberville, E., Le Gall, F., Guilloux, L., Baudoux, A.C., de Vargas, C., Not, F., Thiebaut, E., Henry, N., Simon, N., 2022. Seasonal dynamics of marine protist communities in tidally mixed coastal waters. Mol. Ecol. 31(14), 3761-3783.

Chen, Z.F., Zhang, Q.C., Kong, F.Z., Liu, Y., Zhao, Y., Zhou, Z.X., Geng, H.X., Dai, L., Zhou, M.J., Yu, R.C., 2019. Resolving phytoplankton taxa based on high- throughput sequencing during brown tides in the Bohai Sea. China. Harmful Algae 84, 127–138.

Choi, J., and Park, J.S., 2020. Comparative analyses of the V4 and V9 regions of 18S rDNA for the extant eukaryotic community using the Illumina platform. Scientific Reports 10(1), 1-11.

521 Cipolletta, F., Sarno, D., Morntresor M. Arace, O., Giovanna Bonomo, M., Rossi, R., Soprano,
522 V., Zingone, A., 2019. Potentially toxic microalgae and their toxins in the Gulf of Naples
523 (Mediterranean Sea). Proceedings 19th International Conference on Harmful Algae, Mexico
524 October 2021 (abstract).

525 Cui, Z., Xu, Q., Gibson, K., Liu, S., Chen, N., 2021. Metabarcoding analysis of harmful algal
526 bloom species in the Changjiang Estuary, China. Science of The Total Environment 782,
527 146823.

528 Deagle B.E., L.J. Clarke, J.A. Kitchener, A.M. Polaanowski, A. T. Davidson. 2018. Genetic
529 monitoring of open ocean biodiversity: An evaluation of DNA metabarcoding for processing
530 continuous plankton recorder samples. Molecular Ecology Resources 18, 391-406.

531 del Campo, J., Kolisko, M., Boscaro, V., Santoferrara, L.F., Nenarokov, S., Massana, R.,
532 Guillou, L., Simpson, A., Berney, C., de Vargas, C., Brown, M.W., Keeling, P.J., Wegener
533 Parfrey, L., 2018. EukRef: Phylogenetic curation of ribosomal RNA to enhance understanding of
534 eukaryotic diversity and distribution. PLoS Biol. 16(9), e2005849.

535 de Luca, D., Kooistra, W.H.C.F., Sarno, D., Gaonkar, C.C., and Piredda, R. 2019. Global
536 distribution and diversity of *Chaetoceros* (Bacillariophyta, Mediophyceae): integration of
537 classical and novel strategies. PeerJ 7, e7410.

538 de Luca, D., Piredda, R., Sarno, D., Kooistra, W.H.C.F., 2021. Resolving cryptic species
539 complexes in marine protists: phylogenetic haplotype networks meet global DNA metabarcoding
540 datasets. ISME J 15, 1931–1942.

541 de Vargas, C. , Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le
542 Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.-M. Bittner, L.

Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O. , Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G. Lukeš, J., Malviya, S., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M. Searson, S., Kandels-Lewis, S., T. O. Coordinators, Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., Karsenti, E., Boss, E., Follows, M., Karp-Boss, L., Krzic, U., Reynaud, E.G., Sardet, C., Sullivan, M.B., Velayoudon, D.. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* 348, 1261605.

Emanuel, K., 2017. Assessing the present and future probability of Hurricane Harvey's rainfall. Proc Natl Acad Sci USA 114(48), 12681-12684.

Fiorendino, J.M., Gaonkar, C.C., Henrichs, D.W., Campbell, L., 2021. Drivers of microplankton community assemblage following tropical cyclones. J. Plankton Res. 10.1093/plankt/fbab073

Fu, Z.X., Piumsomboon, A., Punnarak, P., Uttayarnmanee, P., Leaw, C.P., Lim, P.T., Wang, A.J., Gu, H.F., 2021. Diversity and distribution of harmful microalgae in the Gulf of Thailand assessed by DNA metabarcoding. Harmful Algae 106, 102063.

Funaki, H., Gaonkar, C.C., Kataoka, T., Nishimura, T., Tanaka, K., Yanagida, I., Abe, S., Yamaguchi, H., Nagasaki, K., Adachi, M., 2022. Horizontal and vertical distribution of *Gambierdiscus* spp. (Dinophyceae) including novel phylotypes in Japan identified by 18S rDNA metabarcoding. Harmful Algae 111, 102163.

Gaonkar, C.C., Kooistra, W., Lange, C.B., Montresor, M., Sarno, D., 2017. Two new species in the *Chaetoceros socialis* complex (BACILLARIOPHYTA*): C. sporotruncatus* and *C.*

564 *dichatoensis*, and characterization of its relatives *C. radicans* and *C. cinctus*. J. Phycol. 53(4),
565 889-907.

566 Gaonkar, C.C., Piredda, R., Minucci, C., Mann, D.G., Montresor, M., Sarno, D., Kooistra, W.,
567 2018. Annotated 18S and 28S rDNA reference sequences of taxa in the planktonic diatom family
568 Chaetocerotaceae. PLoS One 13(12).

569 Gaonkar, C.C., Piredda, R., Sarno, D., Zingone, A., Montresor, M., Kooistra, W., 2020. Species
570 detection and delineation in the marine planktonic diatoms *Chaetoceros* and *Bacteriastrum*
571 through metabarcoding: making biological sense of haplotype diversity. Environ. Microbiol.
572 22(5), 1917-1929.

573 Gobler, C.J., Doherty, O.M., Hattenrath-Lehmann, T.K., Griffith, A.W., Kang, Y., Litaker, R.W.,
574 2017. Ocean warming since 1982 has expanded the niche of toxic algal blooms in the North
575 Atlantic and North Pacific oceans. Proc. Natl Acad. Sci. USA 114(19), 4975-4980.

576 Gobler, C.J., 2020. Climate change and harmful algal blooms: Insights and perspective. Harmful
577 Algae 91, 101731.

578 Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user
579 interface for sequence alignment and phylogenetic tree building. Molecular Biology and
580 Evolution 27(2), 221-224.

581 Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., de
582 Vargas, C., Decelle, J., del Campo, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Holzmann, M.,
583 Kooistra, W., Lara, E., Le Bescot, N., Logares, R., Mahe, F., Massana, R., Montresor, M.,
584 Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.L., Siano, R., Stoeck, T., Vaulot, D.,
585 Zimmermann, P., Christen, R., 2013. The Protist Ribosomal Reference database (PR2): a catalog

586 of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids
587 Res. 41(D1), D597-D604.

588 Guiry, M.D., 2012. How many species of algae are there? J. Phycol. 48(5), 1057-1063.

589 Hall, N.S., Litaker, R.W., Kenworthy, W.J., Vandersea, M.W., Sunda, W.G., Reid, J.P., Slone,
590 D.H., Butler, S., 2018. Consortial brown tide - picocyanobacteria blooms in Guantanamo Bay,
591 Cuba. Harmful Algae 73, 30-43.

592 Hallegraeff, G.M., Anderson, D.M., Belin, C., Bottein, M.Y.D., Bresnan, E., Chinain, M.,
593 Enevoldsen, H., Iwataki, M., Karlson, B., McKenzie, C.H., Sunesen, I., Pitcher, G.C., Provoost,
594 P., Richardson, A., Schweibold, L., Tester, P.A., Trainer, V.L., Yñiguez, A.T., Zingone, A.,
595 2021a. Perceived global increase in algal blooms is attributable to intensified monitoring and
596 emerging bloom impacts. Communications Earth and Environment 2(1), 117, 101992.

597 Hallegraeff, G., Enevoldsen, H., Zingone, A. 2021b. Global harmful algal bloom status
598 reporting. Harmful Algae 102, 101992.

599 Hallegraeff, G.M., Schweibold, L., Jaffrezic, E., Rhodes, L., MacKenzie, L., Hay, B., Farrell, H.,
600 2021c. Overview of Australian and New Zealand harmful algal species occurrences and their
601 societal impacts in the period 1985 to 2018, including a compilation of historic records. Harmful
602 Algae 102, 101848.

603 Heisler, J., Glibert, P., Burkholder, J., Anderson, D., Cochlan, W., Dennison, W., Gobler, C.,
604 Dortch, Q., Heil, C., Humphries, E., Lewitus, A., Magnien, R., Marshall, H., Sellner, K.,
605 Stockwell, D., Stoecker, D., Suddleson, M., 2008. Eutrophication and harmful algal blooms: A
606 scientific consensus. Harmful Algae 8(1), 3-13.

607 Horiguchi, T., 1995. *Heterocapsa circularisquama* sp. nov. (Peridiniales, Dinophyceae): A new
608 marine dinoflagellate causing mass mortality of bivalves in Japan. Phycol. Res. 43(3), 129-136.

609 Howarth, R.W., 2008. Coastal nitrogen pollution: A review of sources and trends globally and
610 regionally. Harmful Algae 8(1), 14-20.

611 Huang, H., Xu, Q., Gibson, K., Chen, Y., and Chen, N. 2021. Molecular characterization of
612 harmful algal blooms in the Bohai Sea using metabarcoding analysis. Harmful Algae 106,
613 102066.

614 Iwataki, M., 2008. Taxonomy and identification of the armored dinoflagellate genus
615 *Heterocapsa* (Peridiniales, Dinophyceae). Plankton and Benthos Research 3, 135–142.

616 Jamy, M., Foster, R., Barbera, P., Czech, L., Kozlov, A., Stamatakis, A., Bending, G., Hilton, S.,
617 Bass, D., Burki, F., 2020. Long-read metabarcoding of the eukaryotic rDNA operon to
618 phylogenetically and taxonomically resolve environmental diversity. Mol. Ecol. Res, 20(2), 429-
619 443.

620 John, U., Litaker, R.W., Montresor, M., Murray, S., Brosnahan, M.L., Anderson, D.M., 2014.
621 Formal revision of the *Alexandrium tamarense* species complex (Dinophyceae)taxonomy: the
622 introduction of five species with emphasis on molecular-based (rDNA) classification. Protist 165
623 (6), 779–804.

624 Karlson, B., Andersen, P., Arneborg, L., Cembella, A., Eikrem, W., John, U., West, J.J., Klemm,
625 K., Kobos, J., Lehtinen, S., Lundholm, N., Mazur-Marzec, H., Naustvoll, L., Poelman, P., de
626 Rijcke, M., Suikkanen, S., 2021. Harmful algal blooms and their effects in coastal seas of
627 Northern Europe. Harmful Algae 102, 101989.

628 Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment
629 program. Briefings in bioinformatics 9(4), 286-298.

630 Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., Schloss, P.D., 2013. Development of
631 a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on
632 the MiSeq Illumina sequencing platform. Appl. Environ. Microbiol. 79(17), 5112-5120.

633 Liu, S., Gibson, K., Cui, Z., Chen, Y., Sun, X., Chen, N., 2020. Metabarcoding analysis of
634 harmful algal species in Jiaozhou Bay. Harmful Algae 92, 101772.

635 Lundholm, N., Churro, C, Fraga, S., Hoppenrath, M., Iwataki, M., Larsen, J., Mertens, K.,
636 Moestrup, Ø., Zingone, A., (Eds) (2009 onwards). IOC-UNESCO Taxonomic Reference List of
637 Harmful Micro Algae. Accessed at *https://www.marinespecies.org/hab* on 2022-07-06.
638 doi:10.14284/362

639 Lundholm, N., Moestrup, Ø., Kotaki, Y., Hoef-Emden, K., Scholin, C., Miller, P., 2006. Inter-
640 and intraspecific variation of the *Pseudo-nitzschia delicatissima* complex (Bacillariophyceae)
641 illustrated by rRNA probes, morphological data and phylogenetic analyses. J. Phycol. 42:464–
642 81.

643 McKenzie, C.H., Bates, S.S., Martin, J.L., Haigh, N., Howland, K.L., Lewis, N.I., Locke, A.,
644 Pena, A., Poulin, M., Rourke, W.A., Scarratt, M.G., Starr, M., Wells, T., 2021. Three decades of
645 Canadian marine harmful algal events: Phytoplankton and phycotoxins of concern to human and
646 ecosystem health. Harmful Algae 102, 101852.

647 Mozetič, P., Cangini, M., Francé, J., Bastianini, M., Aubry, F. B., Bužančić, M., Cabrini, M.,
648 Cerino, F., Čalić, M., D'Adamo, R., Drakulović, D., Finotto, S., Fornasaro, D., Grilli, F., Kraus,
649 R., Kužat, N., Pfannkuchen, D.M., Gladan, Ž.N, Pompei, M., Rotter, A., Servadei, I., Skejić, S.,

650 2019. Phytoplankton diversity in Adriatic ports: Lessons from the port baseline survey for the
651 management of harmful algal species. Marine Pollution Bulletin 147, 117-132.

652 Nagai, S., Hida, K., Urusizaki, S., Takano, Y., Hongo, Y., Kameda, T., Abe, K., 2016. Massively
653 parallel sequencing-based survey of eukaryotic community structures in Hiroshima Bay and
654 Ishigaki Island. Gene 576, 681–689.

655 Nagai, S., Urusizaki, S., Hongo, Y., Chen, H., and Dzhembekova, N., 2017. An attempt to semi-
656 quantify potentially toxic diatoms of the genus *Pseudo-nitzschia* in Tokyo Bay, Japan by using
657 massively parallel sequencing technology. Plankton Benthos Research 12, 248-258.

658 Nanjappa, D., Audic, S., Romac, S., Kooistra, W.H.C.F., Zingone, A., 2014. Assessment of
659 Species Diversity and Distribution of an Ancient Diatom Lineage Using a DNA Metabarcoding
660 Approach. PLoS ONE 9(8): e103810.

661 Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero,
662 A., Borja, A., Bouchez, A., Cordier, T., Domaizon, I., Feio, M.J., Filipe, A.F., Fornaroli, R.,
663 Graf, W., Herder, J., van der Hoorn, B., Jones, J.I., Sagova-Mareckova, M., Kahlert, M., 2018.
664 The future of biotic indices in the ecogenomic era: Integrating (e) DNA metabarcoding in
665 biological assessment of aquatic ecosystems. Science of the Total Environment 637, 1295-1310.

666 Piredda, R., Tomasino, M.P., D'erchia, A.M., Manzari, C., Pesole, G., Montresor, M., Kooistra,
667 W.H.C.F., Zingone, A., 2017. Diversity and temporal patterns of planktonic protist assemblages
668 at a Mediterranean Long Term Ecological Research site. FEMS Microbiology Ecology, 93(1),
669 10.1093/femsec/fiw200.

Sakamoto, S., Lim, W.A., Lu, D., Dai, X., Orlova, T., Iwataki, M., 2021. Harmful algal blooms and associated fisheries damage in East Asia: Current status and trends in China, Japan, Korea and Russia. Harmful Algae, 102, 101787.

Santoferrara, L., Burki, F., Filker, S., Logares, R., Dunthorn, M., McManus, G.B., 2020. Perspectives from Ten Years of Protist Studies by High-Throughput Metabarcoding. J. Eukaryot. Microbiol. 67(5), 612-622.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. 75(23), 7537.

Slapeta, J., Lopez-Garcia, P., Moreira, D., 2006. Global dispersal and ancient cryptic species in the smallest marine eukaryotes. Mol. Biol. Evol. 23(1), 23-29.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312-1313.

Steidinger, K.A., 2009. Historical perspective on *Karenia brevis* red tide research in the Gulf of Mexico. Harmful Algae 8(4), 549-561.

Thyng, K.M., Hetland, R.D., Socolofsky, S.A., Fernando, N., Turner, E.L. Schoenbaechler, C., 2020. Hurricane Harvey caused unprecedented freshwater inflow to Galveston Bay. Estuaries and Coasts 43(7), 1836-1852.

690 Tragin M., Zingone A., Vaulot D., 2018. Comparison of coastal phytoplankton composition
691 estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic
692 groups and especially Chlorophyta. Environ Microbiol. 20(2):506-520.

693 Wells, M.L., Trainer, V.L., Smayda, T.J., Karlson, B.S.O., Trick, C.G., Kudela, R.M., Ishikawa,
694 A., Bernard, S., Wulff, A., Anderson, D.M., Cochlan, W.P., 2015. Harmful algal blooms and
695 climate change: Learning from the past and present to forecast the future. Harmful Algae 49, 68-
696 93.

697 WoRMS, 2022. World Register of Marine Species at http:// www.marinespecies.org at VLIZ.
698 (Accessed 22 July 2022).

699 Xu X., Yu, Z., Cheng, F., He, L., Cao, X., Song, X., 2017. Molecular diversity and ecological
700 characteristics of the eukaryotic phytoplankton community in the coastal waters of the Bohai
701 Sea, China. Harmful Algae 61, 13-22.

702 Zingone, A., Escalera, L., Aligizaki, K., Fernández-Tejedor, M., Ismael, A., Montresor, M.,
703 Mozetič, P., Taş, S., Totti, C., 2021. Toxic marine microalgae and noxious blooms in the
704 Mediterranean Sea: A contribution to the Global HAB Status Report. Harmful Algae 102,
705 101843.

706 Zingone, A., D'Alelio, D., Mazzocchi, M.G., Montresor, M., Sarno, D., LTER-MC team, 2019.
707 Time series and beyond: multifaceted plankton research at a marine Mediterranean LTER site.
708 Nature Conservation 34, 273-310.

709

710

711 Figure Legends

712

713 Figure 1. Phytoplankton diversity and relative abundance (%) at infrakingdom/class (outer pie)
714 with color-coded labels. The proportion of HAB species within each infrakingdom/class (inner
715 pie) is indicated in red. (a) V4 marker, 835 OTUs (b) V8-V9 marker, 394 OTUs.

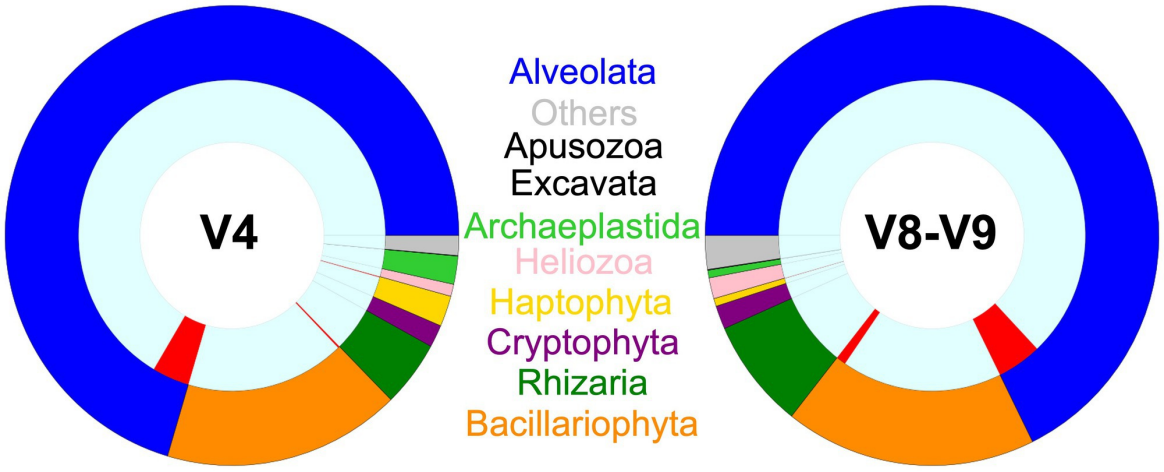
716 Figure 2. Maximum likelihood tree inferred from validated OTUs V4 marker (A) and V8-V9
717 marker (B). Numbers along the internodes represent bootstrap values. Note: in Figure 2B, the
718 long branch is indicated with a gap.

719 Figure 3. Maximum likelihood trees inferred from taxonomically validated HAB references and
720 potential HAB OTUs for the V4 marker (A, C, E, G, I) and V8-V9 marker (B, D, F, H, J).
721 Results for the orders Dinophysiales (A & B), Gonyaulacales (C & D, I & J), Prorocentrales (E
722 & F), Gymnodiniales (G & H). Numbers along the internodes represent bootstrap values.

723

724

725


V4
V8-V9
Alveolata
Others
Apusozoa
Excavata
Archaeplastida
Heliozoa
Haptophyta
Cryptophyta
Rhizaria
Bacillariophyta


Aureococcus anophagefferens U40257
Aureoumbra lagunensis U40258
Vicicitus globosus HQ646558
Pseudochattonella farcimen AM075624
Pseudochattonella verruculosa AM075625
Fibrocapsa japonica AY788931
HRR 28135 11556 115
HRR 23027 9601 437
Heterosigma akashiwo LC214034
Chattonella antiqua AY788921
Chattonella marina HQ710582
Chattonella subsalsa U41649
HRR 27234 6505 274
Nitzschia bizertensis KF955285
HRR 18533 24310 52
Pseudonitzschia calliantha JN091716
Pseudonitzschia pseudodelicatissima GU373965
HRR 20499 21799 1311
Pseudonitzschia caciantha KP708992
Pseudonitzschia granii KJ866907
Pseudonitzschia fraudulenta KX229690
Pseudonitzschia batesiana KP708989
Pseudonitzschia brasiliana KP708990
Pseudonitzschia turgidula FJ222752
Pseudonitzschia kodamae KP709000
Pseudonitzschia lundholmiae KP709002
Pseudonitzschia cuspidata KP708995
HRR 25615 10493 4342
Pseudonitzschia fukuyoi KP708999
Pseudonitzschia galaxiae KJ608079
Pseudonitzschia delicatissima JN091709
Pseudonitzschia seriata GU373969
HRR 17518 11215 17
Pseudonitzschia pungens KP709004
HRR 20630 19330 606
Pseudo-nitzschia multistriata MZ267112
HRR 9868 10560 257
Pseudonitzschia australis GU373961
Pseudonitzschia multiseries AM235382
Chrysochromulina leadbeateri AM491017
Prymnesium polylepis FN551237
Phaeocystis pouchetii AJ278036
Phaeocystis globosa MN927493
HRR 20313 11848 282
Pfiesteria shumwayae AF080098
Pfiesteria piscicida AF330620
Amphidoma languida L5974149
Karlodinium veneficum JN986577
HRR 12065 6847 20140
Heterocapsa circularisquama LC054932v
Protoceratium reticulatum AB727656
Margalefidinium polykrikoides JQ616826
HRR 18833 11457 158
Gonyaulax spinifera FR865625
HRR 7141 10145 332
Polykrikos hartmannii AY421789
HRR 23688 6661 132
Linguladinium polyedra JQ616824
Pyrodinium bahamense DQ500120
HRR 23946 6953 35
Alexandrium minutum AY883006
Alexandrium andersonii JF521619
HRR 3496 7175 45
Alexandrium ostenfeldii LC433644
HRR 11172 16979 1422
Alexandrium tamarense AY883004
Alexandrium pacificum KF908801
Alexandrium catenella JF906989
Alexandrium australiense KF908802
HRR 6965 11458 39
Alexandrium tamiyavanichii AF113935
Alexandrium monilatum AY883005
HRR 18542 11577 858
Alexandrium taylori AJ535385
Alexandrium pseudogoniaulax JF521638
Alexandrium hiranoi AY641564
Gambierdiscus caribaeus HE7750
Gambierdiscus carpenteri KM272970
Gambierdiscus australes AB764308
Gambierdiscus pacificus KY235242
Gambierdiscus balechii KY235239
Gambierdiscus scabrosus AB764271
Gambierdiscus silvae LC661673
Fukuyoa paulensis MW046230
Fukuyoa yasumotoi AB548851
Fukuyoa ruetzleri EF202860
Ostreopsis lenticularis KX129877
Ostreopsis rhodesae KX055855
Ostreopsis ovata AF244939
Prorocentrum concavum Y16237
Prorocentrum lima Y16235
Prorocentrum rhathymum HF565183
Prorocentrum hoffmannianum KF885224
Phalacroma mitra AB551248
Phalacroma rotundatum HM853784
Dinophysis acuminata KJ508017
Dinophysis acuta AJ506973
Dinophysis infundibulus AB366002
Dinophysis norvegica AJ506974
Dinophysis caudata HM853815
Dinophysis fortii AB073118
Dinophysis tripos HM853816
HRR 15234 20623 7
Dinophysis ovum KJ513128
Amphidinium klebsii EU046335
Amphidinium operculatum AB704006
Amphidinium carterae EF492485
Karenia bicuneiformis HM067002
HRR 9264 14041 3
Azadinium poporum HQ324898
HRR 13440 18087 4
Azadinium dexteroporum KR362889
Azadinium spinosum JN680857
Prorocentrum emarginatum Y16239
Gymnodinium catenatum GU362426
HRR 25392 14890 121
Prorocentrum cordatum DQ028763
Prorocentrum mexicanum EF492510
Prorocentrum texanum JQ390504
Karenia papilionacea HM067005
HRR 23600 12937 5
Karenia selliformis HM067007
HRR 6355 7177 88424
Karenia brevis EF492504
Karenia mikimotoi EF492505
0.2


Aureoumbra lagunensis U40258
Aureococcus anophagefferens U40257
Pseudochattonella farcimen AM075624
Pseudochattonella verruculosa AM075625
Vicicitus globosus HQ646558
Nitzschia bizertensis KF938919
Pseudonitzschia pungens U18240
Pseudonitzschia multiseries AM235382
HRR 13898 7503 13892
Pseudonitzschia brasiliana KP708990
Pseudonitzschia lundholmiae KP709002
Pseudonitzschia fraudulenta KX229690
HRR 8840 19569 3
Pseudonitzschia multistriata MZ267112
Pseudo-nitzschia australis GU373961
Pseudonitzschia turgidula FJ222752
Pseudonitzschia seriata GU373969
Pseudonitzschia granii KJ866907
Pseudonitzschia cacianta KP708992
Pseudonitzschia batesiana KP708989
Pseudonitzschia fukuyoi KP708999
Pseudonitzschia calliantha JN091716
Pseudonitzschia galaxiae KJ608079
Pseudonitzschia pseudodelicatissima LC482693
Pseudonitzschia kodamae KP709000
Pseudonitzschia cuspidata KP708995
Pseudonitzschia galaxiae uc GU823527
HRR 23376 12337 8
Pseudonitzschia delicatissima JN091709
Fibrocapsa japonica AY788931
Heterosigma akashiwo LC214034
HRR 22164 2504 36
Chattonella subsalsa U41649
Chattonella marina HQ710582
Chattonella antiqua AY788921
Prymnesium polylepis FN551237
HRR 15562 20696 14
Chrysochromulina leadbeateri AM491017
Phaeocystis pouchetii AJ278036
Phaeocystis globosa MN927493
Gonyaulax spinifera FR865625
HRR 11989 17897 181
Protoceratium reticulatum AB727656
Gambierdiscus silvae LC661673
Gambierdiscus carpenteri KM272970
Gambierdiscus caribaeus HE775087
Gambierdiscus australes AB764308
Gambierdiscus balechii KY235239
Gambierdiscus pacificus KY235242
Fukuyoa paulensis MW046230
Fukuyoa ruetzleri EF202860
Fukuyoa yasumotoi AB548851
Ostreopsis lenticularis KX129877
Ostreopsis ovata AF244939
Ostreopsis rhodesae KX055855
Pyrodinium bahamense AB936750
Alexandrium minutum AY883006
Alexandrium andersonii JF521621
HRR 22911 21770 154
HRR 27353 8387 2196
Alexandrium ostenfeldii LC433647
Alexandrium tamiyavanichii AF113935
Alexandrium tamarense AY883004
Alexandrium australiense KF908802
Alexandrium catenella JF906989
Alexandrium pacificum KF908801
Alexandrium taylori AJ535385
Alexandrium hiranoi AY641564
Alexandrium pseudogoniaulax JF521638
HRR 12613 7084 6
HRR 20217 4346 22
HRR 11513 11319 1067
Alexandrium monilatum AY883005
HRR 16566 3005 181
Prorocentrum concavum Y16237
Prorocentrum lima Y16235
Prorocentrum hoffmannianum KF885224
HRR 17258 19808 7
Karenia papilionacea HM067005
Karenia brevis EF492504
Karenia selliformis HM067007
Karenia mikimotoi EF492505
Margalefidinium polykrikoides JQ616826
HRR 12988 9987 203
Prorocentrum emarginatum Y16239
Amphidoma languida LS974149
Heterocapsa circularisquama LC054932
Polykrikos hartmannii AY421789
Pfiesteria shumwayae AF080098
Pfiesteria piscicida AF330620
Prorocentrum texanum JQ390504
Prorocentrum rhathymum HF565183
Prorocentrum mexicanum EF492510
Amphidinium klebsii EU046335
Amphidinium carterae EF492485
Amphidinium operculatum AB704006
Lingulodinium polyedra JQ616824
Prorocentrum cordatum DQ028763
HRR 21900 23084 1040
Phalacroma rotundatum HM853784
Phalacroma mitra AB551248
Dinophysis acuta AJ506973
Dinophysis infundibulus AB366002
Dinophysis fortii AB073118
Dinophysis acuminata KJ508017
Dinophysis norvegica AJ506974
Dinophysis caudata HM853815
Dinophysis tripos HM853816
HRR 14155 17816 8
Karlodinium veneficum JN986577
Azadinium spinosum JN680857
Azadinium poporum HQ324898
Azadinium dexteroporum KR362889
Gymnodinium catenatum GU362426
HRR 23958 23322 86
0.2

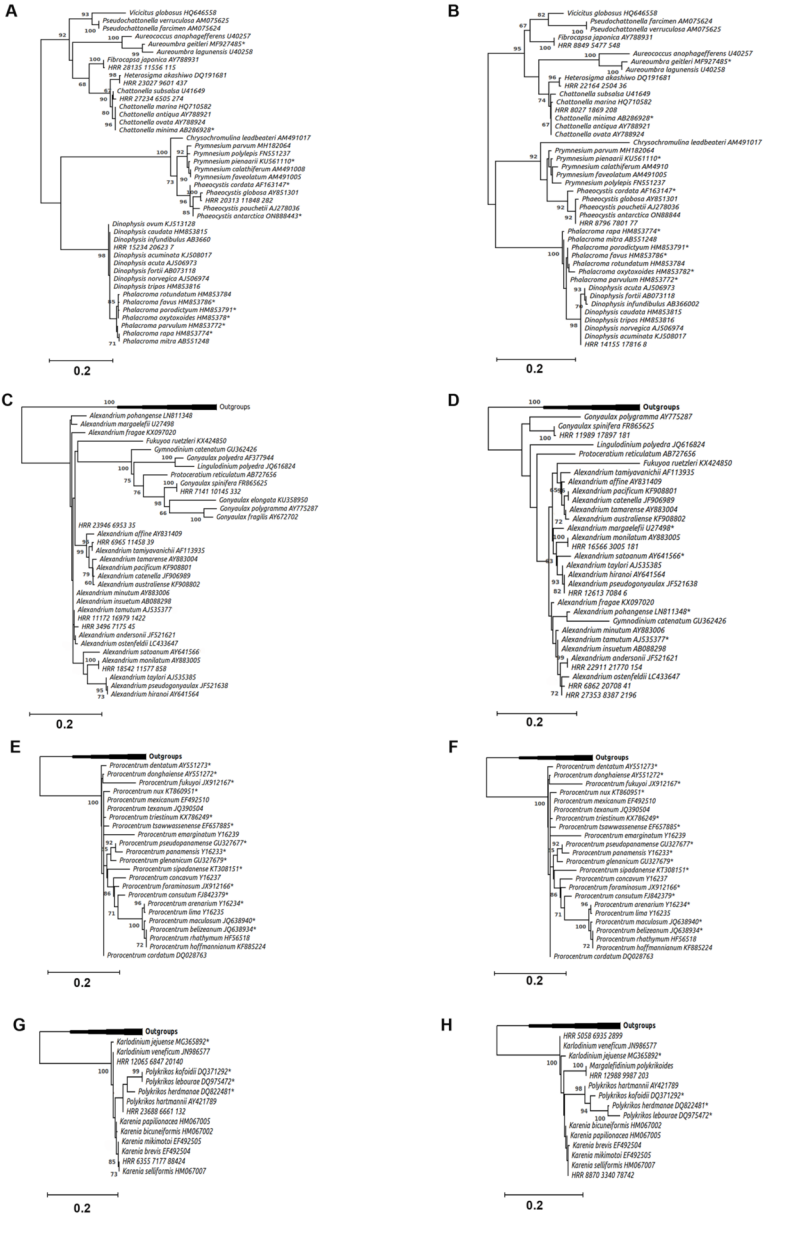

A
Vicicitus globosus HQ646558
Pseudochattonella verruculosa AM075625
Pseudochattonella farcimen AM075624
Aureococcus anophagefferens U40257
Aureoumbra geitleri MF927485*
Aureoumbra lagunensis U40258
Fibrocapsa japonica AY788931
HRR 28135 11556 115
Heterosigma akashiwo DQ191681
HRR 23027 9601 437
Chattonella subsalsa U41649
HRR 27234 6505 274
Chattonella marina HQ710582
Chattonella antiqua AY788921
Chattonella ovata AY788924
Chattonella minima AB286928*
Chrysochromulina leadbeateri AM491017
Prymnesium parvum MH182064
Prymnesium polylepis FN531237
Prymnesium pienaarii KU561110*
Prymnesium calathiferum AM491008
Prymnesium faveolatum AM491005
Phaeocystis cordata AF163147*
Phaeocystis globosa AY851301
HRR 20313 11848 282
Phaeocystis pouchetii AJ278036
Phaeocystis antarctica ON888443*
Dinophysis ovum KJ513128
Dinophysis caudata HM853815
Dinophysis infundibulus AB3660
HRR 15234 20623 7
Dinophysis acuminata KJ508017
Dinophysis acuta AJ506973
Dinophysis fortii AB073118
Dinophysis norvegica AJ506974
Dinophysis tripos HM853816
Phalacroma rotundatum HM853784
Phalacroma fovus HM853786*
Phalacroma parodictyum HM853791*
Phalacroma oxytoxoides HM85378*
Phalacroma parvulum HM853772*
Phalacroma rapa HM853774*
Phalacroma mitra AB551248
0.2
B
Vicicitus globosus HQ646558
Pseudochattonella farcimen AM075624
Pseudochattonella verruculosa AM075625
Fibrocapsa japonica AY788931
HRR 8849 5477 548
Aureococcus anophagefferens U40257
Aureoumbra geitleri MF927485*
Aureoumbra lagunensis U40258
Heterosigma akashiwo DQ191681
HRR 22164 25094 36
Chattonella subsalsa U41649
Chattonella marina HQ710582
HRR 8027 1869 208
Chattonella minima AB286928*
Chattonella antiqua AY788921
Chattonella ovata AY788924
Chrysochromulina leadbeateri AM491017
Prymnesium parvum MH182064
Prymnesium pienaarii KU561110*
Prymnesium calathiferum AM4910
Prymnesium faveolatum AM491005
Prymnesium polylepis FN531237
Phaeocystis cordata AF163147*
Phaeocystis globosa AY851301
Phaeocystis pouchetii AJ278036
Phaeocystis antarctica ON88444
HRR 8796 7801 17
Phalacroma rapa HM853774*
Phalacroma mitra AB551248
Phalacroma parodictyum HM853791*
Phalacroma fovus HM853786*
Phalacroma rotundatum HM853784
Phalacroma oxytoxoides HM853782*
Phalacroma parvulum HM853772*
Dinophysis ovum KJ506973
Dinophysis fortii AB073118
Dinophysis infundibulus AB366002
Dinophysis caudata HM853815
Dinophysis tripos HM853816
Dinophysis norvegica AJ506974
Dinophysis acuminata KJ508017
HRR 14155 17816 8
0.2
C
Outgroups
Alexandrium pohangense LN811348
Alexandrium margalefii U27498
Alexandrium fragae KX097020
Fukuyoa ruetzleri KX424850
Gymnodinium catenatum GU362426
Gonyaulax polyedra AF377944
Lingulodinium polyedra JQ616824
Protoceratium reticulatum AB727656
Gonyaulax spinifera FR865625
HRR 7141 10145 332
Gonyaulax elongata KU358950
Gonyaulax polygramma AY775287
Gonyaulax fragilis AY672702
HRR 23946 6953 35
Alexandrium affine AY831409
HRR 6965 11458 39
Alexandrium tamiyavanichii AF113935
Alexandrium tamarense AY883004
Alexandrium pacificum KF908801
Alexandrium catenella JF906989
Alexandrium australiense KF908802
Alexandrium minutum AY883006
Alexandrium insuetum AB088298
Alexandrium tamutum AJ535377
HRR 11172 16979 1422
HRR 3496 7175 45
Alexandrium andersonii JF521621
Alexandrium ostenfeldii LC433647
Alexandrium satoanum AY641566
Alexandrium monilatum AY883005
HRR 18542 11577 858
Alexandrium taylori AJ535385
Alexandrium pseudogonyaulax JF521638
Alexandrium hiranoi AY641564
0.2
D
Outgroups
Gonyaulax polygramma AY775287
Gonyaulax spinifera FR865625
HRR 11989 17897 181
Lingulodinium polyedra JQ616824
Protoceratium reticulatum AB727656
Fukuyoa ruetzleri KX424850
Alexandrium tamiyavanichii AF113935
Alexandrium affine AY831409
Alexandrium pacificum KF908801
Alexandrium catenella JF906989
Alexandrium tamarense AY883004
Alexandrium australiense KF908802
Alexandrium margalefii U27498*
Alexandrium monilatum AY883005
HRR 16566 3005 181
Alexandrium satoanum AY641566*
Alexandrium taylori AJ535385
Alexandrium hiranoi AY641564
Alexandrium pseudogonyaulax JF521638
HRR 12613 7084 6
Gymnodinium catenatum KX097020
Alexandrium pohangense LN811348*
Gymnodinium catenatum GU362426
Alexandrium minutum AY883006
Alexandrium tamutum AJ535377*
Alexandrium insuetum AB088298
Alexandrium andersonii JF521621
HRR 22911 21770 154
Alexandrium ostenfeldii LC433647
HRR 6862 20708 41
HRR 27353 8387 2196
0.2
E
Outgroups
Protoceratium dentatum AY551273*
Protoceratium donghaiense AY551272*
Protoceratium fukuyoi JX912167*
Protoceratium nux KT860951*
Protoceratium mexicanum EF492510
Protoceratium texanum JQ390504
Protoceratium tristinum KX786249*
Protoceratium tsawwassenense EF657885*
Protoceratium emarginatum Y16239
Protoceratium pseudopanamense GU327677*
Protoceratium panamensis Y16233*
Protoceratium glenanicum GU327679*
Protoceratium sipadanense KT308151*
Protoceratium concavum Y16237
Protoceratium foraminosum JX912166*
Protoceratium consutum FJ842379*
Protoceratium arenarium Y16234*
Protoceratium lima Y16235
Protoceratium maculosum JQ638940*
Protoceratium belizeanum JQ638934*
Protoceratium rhathymum HF56318
Protoceratium hoffmannianum KF885224
Protoceratium cerdotum DQ028763
0.2
F
Outgroups
Protoceratium dentatum AY551273*
Protoceratium donghaiense AY551272*
Protoceratium fukuyoi JX912167*
Protoceratium nux KT860951*
Protoceratium mexicanum EF492510
Protoceratium texanum JQ390504
Protoceratium tristinum KX786249*
Protoceratium tsawwassenense EF657885*
Protoceratium emarginatum Y16239
Protoceratium pseudopanamense GU327677*
Protoceratium panamensis Y16233*
Protoceratium glenanicum GU327679*
Protoceratium sipadanense KT308151*
Protoceratium concavum Y16237
Protoceratium foraminosum JX912166*
Protoceratium consutum FJ842379*
Protoceratium arenarium Y16234*
Protoceratium lima Y16235
Protoceratium maculosum JQ638940*
Protoceratium belizeanum JQ638934*
Protoceratium rhathymum HF56318
Protoceratium hoffmannianum KF885224
Protoceratium cerdotum DQ028763
0.2
G
Outgroups
Karlodinium jejuense MG365892*
Karlodinium veneficum JN986577
HRR 12065 6847 20140
Polykrikos kofoidii DQ371292*
Polykrikos lebourae DQ975472*
Polykrikos herdmanae DQ822481*
Polykrikos hartmannii AY421789
HRR 23688 6661 132
Karenia papilionacea HM067005
Karenia bicuneiformis HM067002
Karenia mikimotoi EF492505
Karenia brevis EF492504
HRR 6355 7177 88424
Karenia selliformis HM067007
0.2
H
Outgroups
HRR 5058 6935 2899
Karlodinium veneficum JN986577
Karlodinium jejuense MG365892*
Margalefidinium polykrikoides
HRR 12988 9987 203
Polykrikos hartmannii AY421789
Polykrikos kofoidii DQ371292*
Polykrikos herdmanae DQ822481*
Polykrikos lebourae DQ975472*
Karenia bicuneiformis HM067002
Karenia papilionacea HM067005
Karenia brevis EF492504
Karenia mikimotoi EF492505
Karenia selliformis HM067007
HRR 8870 3340 78742
0.2

Table 1. HAB species in the Gulf of Mexico. + indicates species present in GoM

	TOAST[3,4]	Anderson et al. 2021	Zingone et al. 2021	Cipolletta et al. 2021	Bresnan et al. 2021	Hallegraeff et al. 2021c	Karlson et al. 2021	Sakamoto et al. 2021	Mozetič et al. 2019	Al-Azari et al. 2015	McKenzie et al. 2021	Xu et al. 2017	Huang et al. 2021	Cui et al. 2021	Fu et al. 2021
Location	Texas coast Gulf of Mexico	Gulf of Mexico	Mediterranean Sea	LTER-MC	Atlantic coast[5]	Australian and New Zealand	Coastal seas of N. Europe[6]	China, Japan, Korea and Russia	Adriatic Ports	Oman /Arabian Sea	British Columbia, Canada	Bohai Sea	Bohai Sea	Changjiang estuary	Gulf of Thailand
Sampling period	2007 - 2021	1990 - 2019		1984 - 2021	1989 - 2019	1985 - 2018	1987 - 2019	1970 - 2018	2011, 2014, 2015	Jun 2006-Apr 2011	1988 - 2017	Apr-Oct 2013	Jul- Aug 2019	Mar 8 – 12, 2019	July - Dec 2018
Number of total HAB species[1]	46	98	84	54	80	56 & 64	51	80	52	24	47	71	74	86	17
Toxic HAB species[2]	46	84	84	54	74	43	51	80	36	12	35	10	36	32	17
HAB species in common[3]	36	31	26	22	23	13	17	19	13	5	19	5	16	13	5
Alexandrium andersonii	+	+	+	+	+			+			+		+	+	
Alexandrium hiranoi[*7]	+														
Alexandrium minutum[#]	+	+	+	+	+	+	+	+	+			+	+		
Alexandrium monilatum	+	+													
Alexandrium ostenfeldii	+	+	+	+	+	+	+	+			+		+	+	
Alexandrium pacificum[*7]	+		+			+	+	+						+	
Alexandrium tamiyavanichii	+	+													
Amphidoma languida[7]	+				+		+		+				+		
Azadinium poporum[#]	+	+			+			+						+	+
Chattonella subsalsa[#]	+	+	+	+							+		+		
Dinophysis norvegica	+	+			+		+	+		+	+		+	+	
Fibrocapsa japonica	+	+	+	+	+			+	+			+	+	+	
Gonyaulax spinifera	+	+	+	+	+		+		+		+	+			
Heterosigma akashiwo	+	+	+	+		+	+	+	+		+	+	+		
Karenia bicuneiformis[#]	+	+	+	+					+						
*Karenia brevis**	+	+	+	+	+	+				+			+	+	
Karenia mikimotoi	+	+	+	+	+	+	+	+	+		+		+	+	+
Karenia papilionacea	+	+	+	+	+				+				+	+	
Karenia selliformis[#]	+	+	+	+											+
Karlodinium veneficum	+	+	+	+	+	+	+	+			+		+	+	
Margalefidinium polykrikoides	+	+	+	+	+			+			+	+			+
Nitzschia bizertensis[#7]	+		+												
Phaeocystis globosa	+	+	+	+			+	+							
Phaeocystis pouchetii	+	+					+				+				
Polykrikos hartmannii[#]	+	+	+	+	+										
Prorocentrum cordatum	+	+	+	+	+				+		+		+	+	
Prorocentrum lima	+	+	+	+	+		+	+	+		+				
Prymnesium polylepis[7]	+		+	+	+		+				+				
*Pseudo-nitzschia australis**	+	+	+		+	+					+		+		
Pseudo-nitzschia cuspidata	+	+	+		+	+	+	+							
*Pseudo-nitzschia delicatissima**	+	+	+	+	+	+	+	+	+	+	+		+	+	+
Pseudo-nitzschia multiseries	+	+	+		+	+	+	+			+		+		
Pseudo-nitzschia pungens	+	+	+	+	+	+	+	+	+	+	+			+	
Pseudo-nitzschia seriata	+	+			+		+	+	+	+	+				

Pseudo-nitzschia turgidula	+	+		+		+		+			+				
Pyrodinium bahamense	+	+													

[1] author identified HAB species (includes high biomass bloom forming non-toxic species)

[2] UNESCO-IOC defined toxic HAB species

[3] this study

[4] includes additional HAB species observed from images at TOAST (*Chattonella sp., Coolia tropicalis, Dinophysis caudata, D. ovum, Gymnodinium catenatum, Lingulodinium polydra, Ostreopsis ovata, Phalacroma rotundatum, Prorocentrum texanum, Pseudochattonella sp., Pseudo-nitzschia multistriata, Vicicitus globosus*)

[5] Coast included Spain, Portugal, France, UK, Ireland, Iceland and Faroe Islands

[6] coasts of the Baltic Sea, Kattegat-Skagerrak, eastern North Sea, Norwegian Sea, and the Barents Sea

[7] first documented in the Gulf of Mexico

* indicates species distinguished with V8-V9 marker only

[#] indicates species distinguished with V4 marker only