Single-cell PCR of the luciferase conserved catalytic domain reveals a unique cluster in the toxic bioluminescent dinoflagellate *Pyrodinium bahamense*

Kathleen D. Cusick^{1,6,*}, Steven W. Wilhelm^{1,2}, Paul E. Hargraves^{4,5},
Gary S. Sayler^{1,2,3,4,5}

¹Center for Environmental Biotechnology, The University of Tennessee, 676 Dabney Hall, Knoxville, TN 37996, USA

²Department of Microbiology, The University of Tennessee, Knoxville, TN 37996, USA

³Department of Ecology and Evolutionary Biology, The University of Tennessee, Knoxville, TN 37996, USA

⁴Harbor Branch Oceanographic Institute of Florida Atlantic University, Ft Pierce, FL 34946, USA

⁵UT-ORNL Joint Institute of Biological Sciences, Oak Ridge, TN 37831, USA

⁶Present address: US Naval Research Lab, Washington, DC 20375, USA

ABSTRACT: *Pyrodinium bahamense* is a toxic, bioluminescent dinoflagellate with a record of intense bloom formation in both the Atlantic-Caribbean and Indo-Pacific regions. To date, limited genetic information exists for *P. bahamense* in comparison to other closely related harmful algal bloom taxa such as *Alexandrium*, or other bioluminescent taxa such as *Pyrocystis*. This study utilized single-cell PCR to explore the molecular diversity of *P. bahamense* within the Indian River Lagoon (IRL), Florida, USA, and a bioluminescent bay in Puerto Rico. *Pyrodinium*-specific primers targeting a ca.1.2-kb region of the 18S rRNA gene and degenerate primers targeting the conserved catalytic domain of the luciferase gene (*lcf*) were applied to single cells isolated from both geographic regions as well as single cells of clonal isolates from the IRL. Phylogenetic analysis revealed that while *P. bahamense* is more closely related to *Alexandrium* spp. at the 18S rRNA gene level, its *lcf* sequences are more closely related to *Pyrocystis* spp. than *Alexandrium* spp. *Pyrodinium bahamense lcf* sequences from the Western Atlantic formed 2 distinct clusters. These clusters were defined by a set of core amino acid substitutions, and the extent of variation was greater than that recorded between the established variants of *Pyrocystis lcf*. *lcf* sequences from an Indo-Pacific strain formed a third distinct cluster. Based on these results, the potential of *lcf* for use in tracking sub-populations of *P. bahamense* is discussed.

KEY WORDS: Indian River Lagoon · Dinoflagellate · *Pyrodinium bahamense* · Saxitoxin · Single-cell PCR · Luciferase · Bioluminescence

INTRODUCTION

Harmful algal blooms (HABs) have increased in both frequency and geographical distribution over the past several decades to the point that every coastal nation is now affected (Wang et al. 2008). *Pyrodinium bahamense* is a toxic HAB species with a

record of repeated intense blooms in the Indian River Lagoon (IRL), Florida, USA, as well as various locations throughout the Indo-Pacific (Phlips et al. 2011, Usup et al. 2012). However, information pertaining to its genetic diversity is lacking relative to other HAB species. *Pyrodinium bahamense* was previously designated as a single species with 2 varieties based on

*Corresponding author: kdcusick@gmail.com

morphological and biochemical differences (Steidinger et al. 1980). The Indo-Pacific variety was designated '*compressum*', and the Atlantic-Caribbean variety '*bahamense*'. For many years, the 2 varieties were believed to be geographically segregated, though reports now exist in which the range of both extends beyond these original locations (Martinez-Lopez et al. 2007, Morquecho 2008), and in fact, these varieties co-occur in several areas (Glibert et al. 2002, Gárate-Lizárraga & González-Armas 2011). In addition to several morphological attributes, one of the primary differences between the 2 was the perceived absence of toxin production in var. *bahamense* (Steidinger et al. 1980). However, *P. bahamense* was defined as the source of saxitoxin found within the tissues of some fish species in the IRL (Quilliam et al. 2004, Landsberg et al. 2006), marking the first recorded occurrences in the Western Atlantic. A recent re-evaluation found no consistent morphological traits that could be used to separate the varieties, while phylogenetic analysis based on large subunit rRNA gene sequences demonstrated the existence of Indo-Pacific and Atlantic-Caribbean ribotypes. This has led to the suggestions that *P. bahamense* is a species complex and that the use of 'varietal' designations is no longer warranted (Mertens et al. 2015).

One notable attribute of *P. bahamense* is that it is bioluminescent (Biggley et al. 1969). The biological production of light is widespread in the marine environment. Bioluminescent species occur in many of the major marine phyla ranging from bacteria to fish, with dinoflagellates serving as the primary source of bioluminescence in surface waters. In general, the reaction entails the oxidation of a substrate, termed luciferin, by molecular oxygen via the luciferase enzyme. In dinoflagellates, bioluminescence is localized to organelles known as scintillons (DeSa & Hastings 1968), which contain the luciferin and luciferase, and in some genera, a luciferin-binding protein (Schmitter et al. 1976, Johnson et al. 1985, Nicolas et al. 1991). Sequencing of the full-length luciferase gene (*lcf*) gene in 7 photosynthetic species of dinoflagellates (*Lingulodinium polyedrum*, *Pyrocystis lunula*, *P. fusiformis*, *P. noctiluca*, *Alexandrium affine*, *A. tamarense*, and *Protoceratium reticulatum*) revealed a unique structure shared among all 7 species: *lcf* encodes a single polypeptide consisting of an N-terminal region of unknown function followed by 3 homologous domains, each with catalytic activity (Okamoto et al. 2001, Liu et al. 2004). Individual domains among species were found to be more similar than among the 3 different domains of the same

species (i.e. D1 of *P. noctiluca* is more similar to D1 of *P. lunula* than to D2 of *P. noctiluca*) (Liu et al. 2004). The gene occurs in multiple copies with tandem organization (Okamoto et al. 2001), and intra- and intermolecular pairwise comparisons between *P. lunula* and *L. polyedrum* support a hypothesis that the *lcf* domains originated through duplication events that occurred prior to the divergence of these species (Okamoto et al. 2001). In contrast to the luciferase of photosynthetic species, the luciferase of *Noctiluca scintillans*, a large, heterotrophic dinoflagellate that is phylogenetically distant (based on 18S rRNA gene sequences) from the 7 bioluminescent species studied previously, differs greatly. The structural organization of *N. scintillans* is composed of 2 distinct domains within the single polypeptide. One domain, located near the N terminus, is similar in sequence to the individual luciferase domains of photosynthetic species, while a domain situated near the C terminus possesses sequence similarity to the luciferin-binding protein (gene: *lbp*) of *L. polyedrum* (Liu & Hastings 2007). Recently, a suite of degenerate primers was developed for *lcf* and used to assess its distribution and genetic diversity in numerous dinoflagellate species (Valiadi et al. 2012), though *P. bahamense* was not among the species studied.

Due to its suggested role in laboratory experiments as a deterrent to grazing (Esaías & Curl 1972, Buskey & Swift 1983), it has been postulated that bioluminescence in dinoflagellates may play a significant role in bloom formation and overall ecosystem dynamics (Valiadi et al. 2012). Additionally, there exists within single dinoflagellate species both bioluminescent and non-bioluminescent strains (for a comprehensive, updated list, the reader is directed to Marcinko et al. 2013). In general, bioluminescence is believed to function as a survival strategy (Hackett et al. 2004). While the experimental evidence to date suggests that bioluminescence acts as a 'burglar alarm' (Burkenroad 1943), in which the flashes alert visual predators (such as fish) to the presence of the dinoflagellate grazers (such as zooplankton) and thus might help maintain the bloom by causing increased predation on grazers, there is still the question of whether bioluminescence plays any role in initiating the bloom.

Based on its bioluminescence potential and bloom-forming tendencies, the primary objective of this study was to examine the molecular diversity of *P. bahamense* in the Western Atlantic based on 18S rRNA and *lcf* sequences. Overall, limited genetic information exists for *P. bahamense* from the Western Atlantic. Therefore, single-cell PCR was applied to cells collected from multiple locations in the IRL and a bio-

luminescent bay in Puerto Rico. PCRs utilized *P. bahamense*-specific 18S rRNA gene primers and primers that targeted the conserved catalytic domain of *lcf*. Ribosomal RNA gene sequences have become common in species identification and phylogenetic classifications among dinoflagellates (Lenaers et al. 1991, John et al. 2003, Yamaguchi & Horiguchi 2005), while functional genes, such those coding for the large subunit of RUBISCO, nitrate reductases, and nitrate transporters, have been used to assess the diversity and structure within phytoplankton communities (Bhadury & Ward 2009, Kang et al. 2011). As no publicly available cultures of *P. bahamense* from the Atlantic-Caribbean exist, clonal isolates were also established from single cells collected from the IRL.

MATERIALS AND METHODS

Water samples were collected from multiple locations in the IRL from 2008 to 2013 (BR, BRBD, IR, 520; Fig. 1) and from Mosquito Bay (MB), Puerto Rico, from 2012 to 2013. A complete list of samples, including locations, dates, name, and sequencing information, can be found in Tables S1 & S2 in Supplement 1 at www.int-res.com/articles/suppl/b025p139_supp.pdf. A total volume of 6 l water was pumped from 0.5 m below the surface and concentrated by successive passage through 2 mm and 125 μ m mesh before collection onto a 35 μ m mesh (Sea Gear); the plankton collected onto the 35 μ m mesh were rinsed into a sterile 1 l bottle using the filtrate. Alternatively, grab samples were collected from surface waters without size fractionation. Samples were placed on ice for transport back to the laboratory and subsequent cell isolation. Whole-water grab samples were also collected from Mosquito Bay, where non-toxic *Pyrodinium bahamense* blooms occur on a yearly basis. Samples were processed in the same manner as those collected from the IRL.

Single cells were isolated for lysis and subsequent PCR using the following protocol: approx. 1 ml volumes of sample water were placed in a Petri dish and viewed with light microscopy. *P. bahamense* cells were identified under 400 $\times$ and 1000 $\times$ magnification based on morphological features previously defined in the literature (Tomas 1997). Individual cells were isolated with a sterile glass micropipette, transferred 2–3 times to drops of sterile HPLC water to facilitate removal of contaminants, and placed in thin-walled 200 μ l sterile PCR tubes. The final volume of water in each tube was brought to 10 μ l with HPLC-purified water. Samples were stored at -20°C until DNA

Fig. 1. Map of sampling sites in the Indian River Lagoon, FL, USA, from which *Pyrodinium bahamense* cells were collected from 2008 to 2013

extraction. DNA extraction consisted of 5 consecutive freeze–thaw cycles alternating between baths of a dry ice/ethanol slurry and heating to 100°C . PCR reagents were added directly to the tubes and the reactions were performed using primers targeting the 18S rRNA gene or *lcf* (i.e. a single PCR was performed per tube).

To obtain 18S rRNA gene sequences, *Pyrodinium*-specific primers were designed based on published *P. bahamense* DNA sequences. The nearly full-length *P. bahamense* sequences listed in the NCBI database were aligned using ClustalX and primers were designed based on regions of sequence identity. The resulting primer set Pcomp370F/Pcomp1530R (Table 1) amplified a ca. 1.2 kb region. PCRs consisted of 1 $\times$ buffer, 200 μ M each dNTP, 1.5 mM MgCl_2 , 200 nm each forward and reverse primer, 1 U Platinum *Taq* DNA polymerase (Invitrogen), template (lysed cell), and were adjusted to a final volume of 50 μ l with nuclease-free water. Samples were subjected to a modified touchdown PCR program consisting of an initial denaturation at 95°C for 3 min/ 94°C for 2 min; 3 cycles each of 94°C for 30 s,

Table 1. List of primers used in the study

Primer name	Sequence (5'–3')	Source
Pcomp370F	AAATTACCCAATCCTGACACT	Present study
Pcomp1530R	CTGATGACTCAGGCTTACT	Present study
DinoLcfF4	CGGCTACGTGCCAARACNAAYCC	Valiadi et al. (2012)
DinoLcfR1	TGCCGGACTCCATCTCCCARAARAA	Valiadi et al. (2012)

65°C/63°C/61°C for 30 s, and 72°C for 1 m 15 s; 33 cycles in which the annealing temperature was decreased to 52°C; and a final extension of 72°C for 10 min. Commercial cultures of *P. bahamense* from which to extract DNA for optimization were not available. Therefore, genomic DNA extracted from pure cultures of *Karodinium veneficum*, a toxic dinoflagellate closely related to *P. bahamense* and also found in the IRL (Phlips et al. 2011), the unicellular green alga *Chlamydomonas reinhardtii*, and *Phaeobacter* (Roseobacter) sp. Y4I, a bacterial genus known to associate with dinoflagellates (Miller & Belas 2006), served as negative controls from which to optimize reaction specificity. The overall PCR success rate was 51 % (SD = 29 %). This included a total of 63 samples. From single cells collected from the IRL, the success rate was 57 % (SD = 20 %) for 36 samples. The success rate was 41 % for single cells collected from Mosquito Bay (SD = 40 %) on a total of 27 samples.

PCRs targeting *lcf* in single cells of *P. bahamense* utilized the degenerate primer pair DinoLcfF4/DinoLcfR1 (Table 1), which amplified a ca. 370 bp region encoding the conserved catalytic domain. PCRs consisted of 1× buffer, 250 µM each dNTP, 1.5 mM MgCl₂, 200 nm each forward and reverse primer, 1.25 U GoTaq Flexi polymerase (Promega), and template, and were brought to a final volume of 50 µl with nuclease-free water. Samples were subjected to a touchdown PCR consisting of an initial denaturation of 95°C for 5 min; 19 cycles of 95°C for 30s, 61°C for 30 s, and 68°C for 1 min, with the annealing temperature decreased 0.5°C every cycle; 30 cycles in which the annealing temperature was maintained at 52°C; and a final extension at 68°C for 5 min. A culture of *Pyrocystis lunula* was kindly provided by Dr. Edith Widder, Ocean Research and Conservation Association, Ft. Pierce, Florida and maintained in f/2 medium at 20°C until cell isolation. This species is known to possess multiple *lcf* variants, and so was used to confirm (1) that the primers amplified the catalytic domain of all *lcf* variants within a cell and (2) that these variations were evident in the resulting sequencing chromatograms. Single cells were

isolated and washed as described for *P. bahamense*. DNA extraction, PCRs using the degenerate primer pair LcfF4/R1, and cloning (when applicable) were performed as for *P. bahamense*.

To examine molecular diversity within cells and among populations, *P. bahamense* clonal isolate cultures

were established from various locations in the IRL. Whole-water samples were collected and maintained at ambient temperature in the dark for transport to the laboratory. Upon arrival, the samples were filtered through 35 µm mesh. The filtrate collected on the mesh was rinsed with sterile seawater into Petri dishes for immediate sorting. The filtered water was then passed via vacuum filtration through 0.22 µm filters and microwave sterilized. This water served as the base for the L1 medium (Guillard & Hargraves 1993), so that cells isolated from various locations were placed into the medium containing the water from which they were collected as base. Single cells were washed several times in the appropriate L1 medium, and placed into 24-well polystyrene plates. Approximately 2 ml medium was added to each well. Plates were sealed to minimize media evaporation and maintained in environmental chambers at 26°C, with 24 h illumination from above. Subsequent PCR analysis was performed on single cells or chains of up to 4 cells. The overall success rate for the *lcf* PCR was 38.5 % (SD = 27 %), on a total of 187 samples. The success rate was 50 % (SD = 31 %) for single cells collected from the IRL, on a total of 46 samples; 27 % (SD = 14 %) on 62 samples from Mosquito Bay; and 45 % (SD = 18 %) on 79 samples of cells from clonal isolates from the IRL.

PCR products were either cloned and sequenced from a plasmid or cleaned and sequenced directly. When sequenced from a plasmid, PCR amplicons were cloned into the pCR2.1 vector using a TOPO TA cloning kit (Invitrogen) following the manufacturer's instructions for kanamycin selection, and sequenced using the M13 forward and reverse primers. For direct sequencing, PCR products were cleaned with the QIAquick PCR purification kit (Qiagen) and sequenced with Pcomp370F/Pcomp1530R or DinoLcfF4/R1. In *lcf* PCRs producing multiple bands, the band of the expected size was gel-purified using the QIAquick gel extraction kit prior to sequencing. Prior to sequencing, a portion of the PCRs was screened for the presence of multiple amplicons (such as would be obtained by amplification of different domains or from homologous domains of gene vari-

ants) via restriction digest with the restriction enzyme *Sph*I. Reactions were performed in 20 µl volumes and consisted of 5 U *Sph*I (Promega), 2 µl 10× Buffer K, and 17.5 µl of the PCR product. Reactions were incubated at 37°C for 2 h and were visualized via gel electrophoresis and ethidium bromide staining.

All samples were sequenced on an ABI 3730 genetic analyzer (Applied Biosystems) at the Molecular Biology Resource Facility, University of Tennessee, Knoxville, USA. Sequences were visually inspected and aligned using Sequencher version 4.9. All sequences obtained in this study have been deposited into GenBank with the following accession numbers: KX377133–KX377203.

Sequence data were trimmed of primer and vector sequences and initially evaluated using the BLAST program (Altschul et al. 1990) with comparison against published sequences in GenBank. All subsequent analysis was performed using MEGA (Molecular Evolutionary Genetics Analysis) v6 (Tamura et al. 2013). Phylogenetic trees based on the 18S rRNA gene sequences were inferred using the minimal evolution method. 18S rRNA gene sequences defined as '*P. bahamense*' or '*P. bahamense* var. *compressum*' of sufficient length (>1200 bp) as well as the 18S rRNA gene sequences of species (to the exact strain and clone when possible) used in constructing *lcf* trees were downloaded from GenBank and included in the analysis. The model that best fit the data was the Tamura-Nei model with the assumption that a certain fraction of sites are evolutionary invariable (TN93+I); this model was used to generate the tree. The reliability of the tree was assessed by 1000 bootstrap replications. A neighbor-joining tree produced identical results. DNA sequences from the individual conserved catalytic domains of all dinoflagellate *lcf* of sufficient length as well as all available *P. bahamense lcf* sequences were downloaded from GenBank. The available *P. bahamense lcf* sequences were from transcriptomic studies (GenBank accession nos. PRJNA169246 and PRJNA261859). The transcriptome of one (PRJNA169246) was from that of an active toxic strain (Hackett et al. 2013). The second sequence, though identified as a transcriptomic study, was genomic DNA of a culture designated as '*P. bahamense* var. *compressum*' from Sorsogon Bay, Philippines. No information was provided as to whether the strain was actively producing toxin, yet strains from this area are commonly found to produce toxin. Sequences were aligned with ClustalW and trimmed to equal length. The similarities between the *Pyrodinium* sequences obtained in this study and those from the multiple catalytic domains of other

dinoflagellates were examined based on genetic distance (p-distance), eliminating gaps only in pairwise comparisons (Table S3 in Supplement 2 at www.int-res.com/articles/suppl/b025p139_supp.xlsx). The model that best fit the data was the Kimura 2-parameter model (K2+G) with gamma distributed rates among sites, and was used to generate the trees. Neighbor-joining and minimal evolution trees were constructed with this model, and produced identical results. In an effort to use the longest sequence length possible (i.e. longer than 300 bp), one of the 2 available *P. bahamense lcf* sequences (that from the established toxic strain, PRJNA169246) be excluded. To further examine the phylogenetic relationship of the *Pyrodinium lcf* domains, an analysis was conducted that included this sequence, as well as representative sequences obtained in the present study resulting in sequences of ca. 214 bp in length. The model that best fit these data was the Tamura 3-parameter with gamma-distributed rates among sites (TN93+G). A neighbor-joining tree was constructed with this model. For all phylogenetic analyses, support for the tree nodes was assessed by the bootstrap method using 1000 iterations.

RESULTS

Both 18S rRNA gene phylogenetic reconstruction methods produced nearly identical results and demonstrated that *Pyrodinium bahamense* 18S rRNA gene sequences grouped most closely to *Alexandrium* spp. (Fig. 2). All *P. bahamense* samples, from multiple locations in the IRL, Puerto Rico, and the Indo-Pacific, clustered together based on the 18S rRNA gene sequences.

Unlike the results obtained with 18S rRNA gene sequences, *P. bahamense lcf* sequences from the IRL and Mosquito Bay formed 2 distinct clusters (Fig. 3). Inclusion of *lcf* sequences from the Indo-Pacific strain resulted in a third cluster (Fig. 3). Thus, to some extent, clusters were defined by geographic location: 2 clusters from the Western Atlantic and one from the Indo-Pacific. While *P. bahamense* grouped most closely to *Alexandrium* spp. based on 18S rRNA gene sequences, *lcf* sequences from *P. bahamense* clustered most closely with those from *Pyrocystis* spp. (Fig. 3). The 2 *P. bahamense* clusters (designated as 'Lcf-WA1' and 'Lcf-WA2' in Figs. 3 & 4) comprising sequences from the Western Atlantic were defined by a set of core amino acid differences (Fig. 4) resulting from non-synonymous substitutions. These differences were not unique to a site or date, as both types of

Fig. 2. Minimal evolution tree of 18S rRNA gene sequences obtained from single cells of *Pyrodinium bahamense* from the Indian River Lagoon (see Fig. 1), Florida, USA, and Mosquito Bay, Puerto Rico. Sequences were retrieved using the *Pyrodinium*-specific primer set Pcomp370F/Pcomp1530R. Published *P. bahamense* 18S rRNA gene sequences and species included in luciferase gene (*lcf*) analysis were included here. The scale bar indicates substitutions over 50% are indicated on branches. Symbols represent sequences obtained from sites as follows: filled circle = site 520, filled diamond = site BR, filled square = site BRBD, see Fig. 1; filled triangle = Mosquito Bay. Dates are indicated as the final set of numbers. Numbers in parentheses indicate clone number from cells in which the PCR product was cloned prior to sequencing

sequences were retrieved over multiple samplings. These sequences were 87% identical (88/101) at the amino acid level. In comparison, the conserved catalytic domain in *P. lunula*, with 3 established *lcf* variants, is 97% identical (98/101 amino acids) between LCFA and LCFB. The genetic distance (p-distance) between the 2 clusters of *P. bahamense* obtained from the Western Atlantic is ca. 0.078. This is greater than the p-distances calculated for any of the homologous domains (using the same sequence length of ca. 300 bp as used for *P. bahamense* p-distance calculations)

of *P. lunula lcfA* and *lcfB*, with distances of 0.049 (D1), 0.033 (D2), and 0.059 (D3) (Table S3 in Supplement 2). In keeping with results obtained from prior studies that found the individual domains of LCFs from different species group together more closely than the different domains of the same species, with the exception of *L. polyedrum*, as evidenced both here and in prior studies (Liu et al. 2004, Valiadi et al. 2012), sequences indicative of domains D1 and D3 from *Pyrodinium* grouped most closely with those of D1 and D3 from *Pyrocystis* spp., respectively (Fig. 5).

Fig. 3. Neighbor-joining tree of luciferase gene (*lcf*) sequences obtained from single cells or clonal isolates of *Pyrodinium bahamense* from the Indian River Lagoon, Florida, USA, and Mosquito Bay, Puerto Rico. Sequences were retrieved using the degenerate primers DinoLcfF4/LcfR1. The evolutionary distances were computed using the Kimura 2-parameter method and are in the units of the number of base substitutions per site. All ambiguous positions were removed for each sequence pair. Sequences had a final nucleotide length of 322 bp. The scale bar indicates substitutions per site. Bootstrap values for 1000 iterations over 50% are indicated on branches. See Fig. 2 for symbols and sites. Dates are indicated as the final 4 numbers (mmyy, all years after 2000) in both single cells from the environment (designated by 'E') and clonal isolates (designated by 'C'). Numbers in parentheses indicate clone number from cells in which the PCR product was cloned prior to sequencing. Open diamonds indicate *lcf* sequences obtained from single cells of a pure culture of *P. lunula* (=PI), with number representative of clone. 'D1', 'D2', or 'D3' following GenBank accession number indicates whether the sequence corresponds to domain 1, 2, or 3. The 2 clusters are designated as 'Lcf-WA1' and 'Lcf-WA2' in reference to the core amino acid differences between the 2 sets of sequences

Lcf-WA1 : HGRWTTVSGGQAAFINKAIEAGTLGAAEAHKIADTDHHQTGGMYVRANQFGDVCTAYAS : 60
 Lcf-WA2 : HGRWTTVSGGQAAFINKAIEAGMLGAAEAHKIADTDHHQTGGMYLRINQFGDVCTVDAS : 60

Lcf-WA1 : VARFARAKPTWKSGRYFYEPVSGGNRLGVWVLPDEYREIG : 101
 Lcf-WA2 : VAKFARAKPTWKSGHYFYEPVSGGNRLGVWVLPDEYRKIG : 101

Fig. 4. Amino acid differences as a result of nucleotide substitutions between the 2 luciferase gene clusters obtained from single cells of *Pyrodinium bahamense*. One representative sample from each cluster, designated as 'Lcf-WA1' and 'Lcf-WA2', is illustrated

In an attempt to determine whether the 2 *lcf* variants obtained from single cells of *P. bahamense* were indicative of gene variants within the same cell (such as with *P. lunula*) or gene variants within a subpopulation, clonal isolate cultures were established from different locations within the IRL. During the dinoflagellate cell cycle, the cells undergo asexual

division to yield 2 genetically identical daughter cells. Therefore, we hypothesized that a clonal isolate culture established from a single cell would be genetically identical and therefore consistently yield 1 of the 2 sequences, but not both. To test this hypothesis, *lcf* was amplified from tubes containing 1–4 cells from clonal isolate cultures. Single-cell *lcf*

Fig. 5. Individual domains of *Pyrodinium bahamense* luciferase gene (*lcf*) coding for the conserved catalytic region are most similar to *Pyrocystis* spp. domains. Neighbor-joining tree constructed with all available *P. bahamense* *lcf* sequences from GenBank (including TSA and Short Read Archives) and representative sequences obtained in the present study. Sequences were aligned to those obtained in the present study and trimmed to an equal length of 214 nucleotides. All ambiguous positions were removed for each sequence pair. The scale bar indicates substitutions per site. Bootstrap values for 1000 iterations over 50% are indicated on branches. Filled triangles represent sequences obtained in the present study using the primer set DinoL-cfF4/LcfR1. PI = *P. lunula* cells from pure culture. Filled diamonds indicate *P. bahamense* *lcf* sequences obtained from GenBank. D1, D2, or D3 following GenBank accession number indicates whether the sequence corresponds to domain 1, 2, or 3

amplification from the unialgal culture of *P. lunula* served as a 'positive control.'

From a unialgal culture of *P. lunula*, the degenerate primers amplified the D3 domain of both *lcfA* and *lcfB* (as indicated by open diamonds in Fig. 3) from each single cell. This was evident in the sequencing chromatograms, which showed heterozygosity at each base (data not shown) where the 2 variants have established sequence differences. Sequencing of PCRs from *P. bahamense* single cells from the environment produced either of the 2 sequences, but not both from a single PCR of a single cell. Two different methods were used to screen for the presence of multiple products, such as would be obtained with *P. lunula* single cells. We predicted that if DNA sequences from more than one domain were amplified, these sequence differences would be manifested by (1) different banding patterns in restriction digests of the *lcf* PCR; and (2) different clone sequences from the same PCR. PCRs were screened via restriction digests with *SphI*, which does not cut within the sequence defining one of the clusters (Lcf-WA1), but would produce bands of ca. 100 and 250 bp in the other (Lcf-WA2). Digests visualized with gel electrophoresis and ethidium bromide staining showed only a single band in PCRs from one cluster, while PCRs representative of the other cluster displayed the expected 100 and 250 bp banding pattern (data not shown), confirming that PCRs produced a single *lcf* amplicon. Additionally, clones from the same PCR of single cells from the environment produced identical or nearly identical sequences (data not shown), with no heterozygosity among bases as seen with *P. lunula*. While each PCR produced a single *lcf*, both types of *lcf* sequences were obtained from all sampling sites (as illustrated on the neighbor-joining tree in Fig. 3).

Both types of sequences were also obtained from multiple clonal isolate cultures of *P. bahamense* established from 2 of the sites (520 and IR). While heterozygosity at the consistent positions occurred in all *P. lunula* single cell PCRs obtained from the same culture, single cells from *P. bahamense* clonal isolates displayed only 1 of the 2 types of sequence variants (i.e. no heterozygosity in chromatograms). However, both variants were obtained from each *P. bahamense* clonal isolate culture from which multiple single-cell PCRs were performed. In most cases, sequencing of PCRs containing 1–4 cells from the clonal isolate culture produced a single sequence, with no heterozygosity among bases in the chromatograms (data not shown). However, a small percentage (ca. 15%) in which the PCR product was

sequenced directly, typically containing 4 cells, yielded sequences with heterozygosity. The heterozygosity consistently occurred among the same bases. Collectively, these data suggest the *P. bahamense lcf* sequence variations represent gene families within the same cell or clonal population, yet at a much greater level than that seen in bioluminescent species with known gene variants (i.e. *Pyrocystis*).

DISCUSSION

Though a major contributor to blooms along the east coast of Florida, including those resulting in saxitoxin production, and the source of the bioluminescence in 'bioluminescent bays' in Puerto Rico, there is limited genetic data for *Pyrodinium bahamense* from the Western Atlantic. Therefore, *Pyrodinium*-specific 18S rRNA gene primers were designed based on conserved regions of identity from all *P. bahamense* sequences available in GenBank, resulting in a primer set that amplified a ca. 1.2 kb region. Phylogenetic analysis demonstrated that *P. bahamense* 18S rRNA gene sequences from the Western Atlantic were most similar to *Alexandrium* spp. These data support morphological and large ribosomal subunit rRNA gene analyses, in which *P. bahamense* strains inclusive of both Indo-Pacific and Atlantic strains were located close to, but independent from, the clade of *Alexandrium* (Leaw et al. 2005, Mertens et al. 2015).

Functional genes have been used to assess the diversity and structure within phytoplankton communities (Bhadury & Ward 2009, Kang et al. 2011). This study utilized primers targeting the conserved catalytic domain of *lcf* to examine the molecular diversity of *P. bahamense* cells from the Western Atlantic with existing *lcf* sequence information from a range of dinoflagellates, including both toxic (i.e. *Lingulodinium* spp., *Alexandrium* spp.) and non-toxic (i.e. *Pyrocystis* spp.) species. The *lcf* primers employed in the present study had been used previously to retrieve sequences from numerous dinoflagellate species (Valiadi et al. 2012, Valiadi et al. 2014); however, *P. bahamense* was not among those for which sequence information was obtained. As demonstrated with single cells from pure culture, these *lcf* primers selectively amplified the third catalytic domain (D3) of both *P. lunula lcf* variants, but not D1 or D2. This may have also been the case with *P. bahamense*—amplification specific to a certain domain—which would indicate that the *P. bahamense lcf* sequence differences represent gene variants. The 2 sequence

clusters ('Lcf-WA1' and 'Lcf-WA2') obtained in the present study were 87 % identical at the amino acid level; in comparison, this same conserved catalytic domain in *P. lunula* is 97 % identical between LcfA and LcfB. Therefore, *lcf* variation within *P. bahamense* occurs at a much greater level than that seen in bioluminescent species with known gene variants (i.e. *Pyrocystis*) and is similar to that of the differences found for the *lbp* of *L. polyedrum*, which occurs as a gene family with 2 variants that share ca. 86 % identity (Valiadi & Iglesias-Rodriguez 2013). However, the only way to accurately determine whether the 2 sequence types obtained in the present study are gene variants—either within an individual cell or between cells of a sub-population—or 2 domains from the same gene is to sequence the entire gene, from both single cells from the environment and single cells of a clonal isolate culture.

The crystal structure of the *L. polyedrum* luciferase D3 has been resolved at 1.8 Å (Schultz et al. 2005) and so was used as a model to examine structural information pertaining to the non-synonymous substitutions between Lcf-WA1 and Lcf-WA2. Based on the *L. polyedrum* model, the region consists of 7 α -helices and 16 β -strands, organized into 2 sub-domains: a regulatory domain defined by a 3-helix bundle at the top followed by a β -barrel below composed of β -strands 5–14. The *P. bahamense* region amplified in this study spanned β -strand 5, α -helices 5 and 6, and β -strands 6–11. All the amino acid substitutions occurring within the modeled region between the 2 *P. bahamense lcf* variants were non-synonymous, and ca. 67 % of these entailed substitutions with different physical properties (i.e. polar to non-polar, polar to basic, etc.). Except for one, all non-synonymous substitutions occurred within a β -strand (5, 7, 8, and 11) or α -helix 5. Additionally, these substitutions were unique to the Lcf-WA1 sequence cluster, as alignment within this region showed these residues to be conserved among Lcf-WA2, *P. lunula* LcfA and LcfB, *L. polyedrum*, *A. tamarensis*, *A. affine*, and the Indo-Pacific strain of *P. bahamense* (see Fig. S1 in Supplement 1).

Natural populations of *P. bahamense* have been shown to exhibit a dual character to their bioluminescence (Biggley et al. 1969). Additionally, the bright-emitter *P. fusiformis* has also been shown to exhibit 2 forms of bioluminescence: a short, intense flash, and a dimmer, prolonged glow (Widder & Case 1981). It may be possible that one of the variants is used for producing a lower-intensity 'glow' while the second variant produces the more intense 'flash.' It remains to be determined the effect on bioluminescence out-

put of these substitutions, and the underlying physiological mechanisms, such as enhanced substrate (luciferin) binding or the kinetics of the reaction.

Based on the tendency of *P. bahamense* for toxin production combined with the relatively new emergence of toxic strains in the Western Atlantic, an important next step is to examine the correlation(s) between *lcf* variant regulation, bioluminescence intensity, and toxin production. Including *lcf* from the Indo-Pacific suggests that the sequences are indicative of sub-populations. The molecular data obtained here demonstrate a level of diversity within the *P. bahamense lcf* clusters that facilitates the development of molecular probes for use in tracking *P. bahamense* blooms. Knowing the genetic diversity within and among blooms is important for understanding both the origin of blooms emerging in new locations and the mechanisms driving the evolution of the bloom-forming species (Van Dolah et al. 2009). Correlating the *lcf* data presented here with those of other functional genes, such as those involved in toxin biosynthesis, will provide a more comprehensive understanding as to *P. bahamense* bloom dynamics and their influence on overall coastal ecosystem dynamics. No data currently exist as to the correlation between bioluminescence intensity and toxin production, and linking the use of gene variants with toxin production may provide valuable information that can then be used to track sub-populations within the same bloom.

Acknowledgements. The authors thank Drs. Wayne Litaker and Mark Vandersea, NOAA, and Dr. Allen Place, University of Maryland, for providing the *Karodinium* DNA, and Dr. Alison Buchan and Mary Hadden, University of Tennessee, for the *Roseobacter* DNA. We greatly appreciate the scientific advice and support of Dr. Michael Latz (UC San Diego). Many thanks to members of the Aquatic Group at the Kennedy Space Center: Doug Scheidt, Karen Holloway-Adkins, Eric Reier, and Russ Lowers, along with Greg Cusick, for water sample collection. K.D.C. was supported by a NASA graduate student fellowship. This is contribution no. 2041 from the Harbor Branch Oceanographic Institute.

LITERATURE CITED

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. *J Mol Biol* 215: 403–410
- Bhadury P, Ward BB (2009) Molecular diversity of marine phytoplankton communities based on key functional genes. *J Phycol* 45:1335–1347
- Biggley WH, Swift E, Buchanan RJ, Seliger HH (1969) Stimulable and spontaneous bioluminescence in marine dinoflagellates, *Pyrodinium bahamense*, *Gonyaulax polyedra*, and *Pyrocystis lunula*. *J Gen Physiol* 54:96–122

- Burkenroad MD (1943) A possible function of bioluminescence. *J Mar Res* 5:161–164
- Buskey EJ, Swift E (1983) Behavioral responses of the coastal copepod *Acartia hudsonica* (Pinhey) to simulated dinoflagellate bioluminescence. *J Exp Mar Biol Ecol* 72: 43–58
- DeSa R, Hastings JW (1968) The characterization of scintillons. Bioluminescent particles from the marine dinoflagellate, *Gonyaulax polyedra*. *J Gen Physiol* 51:105–122
- Esaias WE, Curl HC (1972) Effect of dinoflagellate bioluminescence on copepod ingestion rates. *Limnol Oceanogr* 17:901–905
- Gárate-Lizárraga I, González-Armas R (2011) Occurrence of *Pyrodinium bahamense* var. *compressum* along the southern coast of the Baja California Peninsula. *Mar Pollut Bull* 62:626–630
- Glibert PM, Landsberg JH, Evans JJ, Al-Sarawi MA and others (2002) A fish kill of massive proportion in Kuwait Bay, Arabian Gulf, 2001: the roles of bacterial disease, harmful algae, and eutrophication. *Harmful Algae* 1: 215–231
- Guillard RRL, Hargraves PE (1993) *Stichochrysis immobilis* is a diatom, not a chrysophyte. *Phycologia* 32:234–236
- Hackett JD, Anderson DM, Erdner DL, Bhattacharya D (2004) Dinoflagellates: a remarkable evolutionary experiment. *Am J Bot* 91:1523–1534
- Hackett JD, Wisecarver JH, Brosnahan ML, Kulis DM and others (2013) Evolution of saxitoxin synthesis in cyanobacteria and dinoflagellates. *Mol Biol Evol* 30:70–78
- John U, Fensome RA, Medlin LK (2003) The application of a molecular clock based on molecular sequences and the fossil record to explain biogeographic distributions within the *Alexandrium tamarense* ‘species complex’ (Dinophyceae). *Mol Biol Evol* 20:1015–1027
- Johnson CH, Inoue S, Flint A, Hastings JW (1985) Compartmentalization of algal bioluminescence: autofluorescence of bioluminescent particles in the dinoflagellate *Gonyaulax* as studied with image-intensified video microscopy and flow cytometry. *J Cell Biol* 100:1435–1446
- Kang LK, Wang HF, Chang JS (2011) Diversity of phytoplankton nitrate transporter sequences from isolated single cells and mixed samples from the East China Sea and mRNA quantification. *Appl Environ Microbiol* 77: 122–130
- Landsberg JH, Hall S, Johannessen JN, White KD and others (2006) Saxitoxin puffer fish poisoning in the United States, with the first report of *Pyrodinium bahamense* as the putative toxin source. *Environ Health Perspect* 114: 1502–1507
- Leaw CP, Lim PT, Ng BK, Cheah MY, Ahmad A, Usup G (2005) Phylogenetic analysis of *Alexandrium* species and *Pyrodinium bahamense* (Dinophyceae) based on theca morphology and nuclear ribosomal gene sequence. *Phycologia* 44:550–565
- Lenaers G, Scholin C, Bhaud Y, Saint-Hilaire D, Herzog M (1991) A molecular phylogeny of dinoflagellate protists (Pyrrhophyta) inferred from the sequence of 24S rRNA divergent domains D1 and D8. *J Mol Evol* 32:53–63
- Liu L, Hastings JW (2007) Two different domains of the luciferase gene in the heterotrophic dinoflagellate *Noctiluca scintillans* occur as two separate genes in photosynthetic species. *Proc Natl Acad Sci USA* 104:696–701
- Liu L, Wilson T, Hastings JW (2004) Molecular evolution of dinoflagellate luciferases, enzymes with three catalytic domains in a single polypeptide. *Proc Natl Acad Sci USA* 101:16555–16560
- Marcinko CLJ, Painter SC, Martin AP, Allen JT (2013) A review of the measurement and modelling of dinoflagellate bioluminescence. *Prog Oceanogr* 109:117–129
- Martinez-Lopez A, Ulloa-Perez AE, Escobedo-Urias DC (2007) First record of vegetative cells of *Pyrodinium bahamense* (Gonyaulacales: Goniodomataceae) in the Gulf of California. *Pac Sci* 61:289–293
- Mertens KN, Wolny J, Carbonell-Moore C, Bogus K and others (2015) Taxonomic re-examination of the toxic armored dinoflagellate *Pyrodinium bahamense* Plate 1906: Can morphology or LSU sequencing separate *P. bahamense* var. *compressum* from var. *bahamense*? *Harmful Algae* 41:1–24
- Miller TR, Belas R (2006) Motility is involved in *Silicibacter* sp. TM1040 interaction with dinoflagellates. *Environ Microbiol* 8:1648–1659
- Morquerocho L (2008) Morphology of *Pyrodinium bahamense* Plate (Dinoflagellata) near Isla San José, Gulf of California, Mexico. *Harmful Algae* 7:664–670
- Nicolas MT, Morse D, Bassot JM, Hastings JW (1991) Colocalization of luciferin binding protein and luciferase to the scintillons of *Gonyaulax polyedra* revealed by double immunolabeling after fast-freeze fixation. *Protoplasma* 160:159–166
- Okamoto OK, Liu L, Robertson DL, Hastings JW (2001) Members of a dinoflagellate luciferase gene family differ in synonymous substitution rates. *Biochemistry* 40: 15862–15868
- Phlips EJ, Badylak S, Christman M, Wolny J and others (2011) Scales of temporal and spatial variability in the distribution of harmful algae species in the Indian River Lagoon, Florida, USA. *Harmful Algae* 10:277–290
- Quilliam M, Wechsler D, Marcus S, Ruck B, Wekell M, Hawryluk T (2004) Detection and identification of paralytic shellfish poisoning toxins in Florida pufferfish responsible for incidents of neurologic illness. In: Steidinger K, Landsberg J, Tomas C, Vargo G (eds) *Harmful Algae 2002*. Florida Fish and Wildlife Conservation Commission of UNESCO, St. Petersburg, FL, p 116–118
- Schmitter RE, Njus D, Siulzman FM, Gooch V, Hastings JW (1976) Dinoflagellate bioluminescence: a comparative study of *in vitro* components. *J Cell Physiol* 87:123–134
- Schultz LW, Liu L, Cegielski M, Hastings JW (2005) Crystal structure of a pH-regulated luciferase catalysing the bioluminescent oxidation of an open tetrapyrrole. *Proc Natl Acad Sci USA* 102:1378–1383
- Steidinger KA, Tester LS, Taylor FJR (1980) A redescription of *Pyrodinium bahamense* var. *compressa* (Bohm) stat. nov. from Pacific red tides. *Phycologia* 19:329–337
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Tomas CR (ed) (1997) Identifying marine phytoplankton. Academic Press, San Diego, CA
- Usup G, Ahmada A, Matsuoka K, Lim PT, Leaw CP (2012) Biology, ecology and bloom dynamics of the toxic marine dinoflagellate *Pyrodinium bahamense*. *Harmful Algae* 14:301–312
- Valiadi M, Iglesias-Rodriguez D (2013) Understanding bioluminescence in dinoflagellates—How far have we come? *Microorganisms* 1:3–25
- Valiadi M, Iglesias-Rodriguez MD, Amorim A (2012) Distribution and genetic diversity of the luciferase gene within marine dinoflagellates. *J Phycol* 48:826–836

- Valiadi M, Painter SC, Allen JT, Balch WM, Iglesias-Rodriguez DM (2014) Molecular detection of bioluminescent dinoflagellates in surface waters of the Patagonian Shelf during early Austral summer 2008. PLOS ONE 9: e98849
- Van Dolah FM, Lidie KB, Monroe EA, Bhattacharya D, Campbell L, Doucette GJ, Kamykowski D (2009) The Florida red tide dinoflagellate *Karenia brevis*: New insights into cellular and molecular processes underlying bloom dynamics. Harmful Algae 8:562–572
- Wang RL, Biales A, Bencic D, Lattier D and others (2008) DNA microarray application in ecotoxicology: Experimental design, microarray scanning, and factors affecting transcriptional profiles in a small fish species. Environ Toxicol Chem 27:652–663
- Widder EA, Case JF (1981) Two flash forms in the bioluminescent dinoflagellate, *Pyrocystis fusiformis*. J Comp Physiol 143:43–52
- Yamaguchi A, Horiguchi T (2005) Molecular phylogenetic study of the heterotrophic dinoflagellate genus *Protoperdinium* (Dinophyceae) inferred from small subunit rRNA gene sequences. Phycol Res 53:30–42

Editorial responsibility: Warwick Vincent,
Sainte-Foy, Quebec, Canada

Submitted: July 20, 2015; Accepted: August 16, 2016
Proofs received from author(s): September 18, 2016