

High intraindividual, intraspecific, and interspecific variability in large-subunit ribosomal DNA in the heterotrophic dinoflagellates *Protoperidinium*, *Diplopsalis*, and *Preperidinium* (Dinophyceae)

KRISTIN E. GRIBBLE*† AND DONALD M. ANDERSON

Woods Hole Oceanographic Institution, Biology Department, MS 32, Woods Hole, MA 02543, USA

K.E. GRIBBLE AND D.M. ANDERSON. 2007. High intraindividual, intraspecific, and interspecific variability in large-subunit ribosomal DNA in the heterotrophic dinoflagellates *Protoperidinium*, *Diplopsalis*, and *Preperidinium* (Dinophyceae). *Phycologia* 46: 315–324. DOI: 10.2216/06-68.1

Large-subunit ribosomal DNA (LSU rDNA) gene sequences are frequently used to infer the phylogeny and species identity of organisms. The many copies of the LSU rDNA found in the genome are thought to be kept homogeneous by concerted evolution. In this study, we found high intraindividual, intraspecific and interspecific diversity in the D1-D6 region of the LSU rDNA in eight species of thecate, heterotrophic dinoflagellates using single-cell PCR, cloning, and sequencing of many clones. For each species, the clone library was usually comprised of one highly represented copy and many unique sequences. Sequence differences were primarily characterized by single base pair substitutions, single base pair insertion/deletions (indels), and/or large indels. Phylogenetic analysis of all sequence clones gave strong support for monophyly of the polymorphic copies of each species and recovered the same species tree as an analysis using just one sequence per species. Analysis of LSU rDNA expression in three species by RT-PCR indicated that expressed copies have fewer substitutions and fewer and smaller indels and that 50% or more of the copies are pseudogenes. High intraspecific and intraindividual LSU rDNA sequence variability could lead to inaccurate species phylogenies and overestimation of species diversity in environmental sequencing studies.

KEY WORDS: *Diplopsalis*, Gene family, Heterotrophic dinoflagellates, Intraspecific variability, LSU rDNA, *Preperidinium*, *Protoperidinium*, Pseudogenes, RT-PCR, Sequence polymorphism

INTRODUCTION

Protoperidinium Bergh is a cosmopolitan genus of marine, thecate, heterotrophic dinoflagellates, of which there are more than 200 species (Balech 1974). *Protoperidinium* spp. occupy an important ecological niche as raptorial, and often selective, feeders on large phytoplankton and protistan prey (Gaines & Taylor 1984; Jacobson & Anderson 1986; Naustvoll 2000). Although an estimated 50% of dinoflagellates are heterotrophic or mixotrophic (Gaines & Elbrächter 1987), these groups are relatively underrepresented in molecular phylogenetic studies. Thecate heterotrophic dinoflagellates, in particular, have been poorly sampled.

Large-subunit ribosomal DNA (LSU rDNA) sequences are frequently used to infer the phylogeny of organisms. The LSU rDNA contains both highly variable and highly conserved regions, making it useful for comparisons of organisms from a range of taxonomic levels (Hillis & Dixon 1991). Known to be tandemly arrayed in hundreds to thousands of copies in most organisms, the LSU rDNA is a relatively easy target for amplification. Additionally, the many copies of ribosomal genes within a species are thought to be kept homogeneous by concerted evolution, a process of nonindependent evolution of repetitive DNA sequences leading to greater similarity of sequences in a gene family within a species than between species (Dover 1982). Concerted evolution, generally accept-

ed as a universal phenomenon (Liao 1999), works by various mechanisms of DNA repair and replication, including unequal crossing over among tandemly arranged genes in the same gene family, gene conversion (Arnheim *et al.* 1980), and reciprocal exchange (Aguilera & Klein 1989).

In recent years, however, workers have begun to realize the extent of both intraspecific and intraindividual variability in nuclear ribosomal genes. Examples of intraspecific variability in rDNA have been found in a range of organisms, including in the variable internal transcribed spacer region (ITS) of pinon pine (Gernandt *et al.* 2001) and the dinoflagellate *Oxhyris marina*, the small-subunit (SSU) rDNA of *Skeletonema* spp. diatoms (Alverson & Kolnick 2005), *Acropora* spp. corals (Márquez *et al.* 2003), and the dinoflagellates *Symbiodinium* spp. (Santos *et al.* 2003) and *Alexandrium fundyense* (Scholin *et al.* 1993) and in the LSU rDNA of dinoflagellates including *Alexandrium catenella* (Yeung *et al.* 1996) and *Dinophysis acuminata* (Rehnstam-Holm *et al.* 2002). Examples of intraindividual rDNA variability are fewer but include the ITS region in cactus (Hartmann *et al.* 2001) and serviceberry bushes (Campbell *et al.* 1997) and sequence polymorphism across the entire rDNA cistron in salmonid fish (Reed *et al.* 2000) and in humans (Brownell *et al.* 1983).

Intraspecific diversity in ribosomal genes has important implications for phylogenetic studies. Orthologous genes, which trace their ancestry to a speciation event, must be used for determining evolutionary species relationships. Paralogous genes, often difficult to distinguish from orthologs in large gene families, trace their ancestry to a duplication event. If

* Corresponding author (kgribble@mbl.edu).

† Present address: Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543, USA.

Table 1. Isolation and culturing information for thecate heterotrophic dinoflagellates studied. Food species: Db = *Ditylum brightwellii* (West) Grunow; Ca = *Chaetoceros affinis* Lauder; Lp = *Lingulodinium polyedrum* (Stein) Dodge; n/a = not applicable.

Species	Culture designation	Clonal (C) /unialgal (U) /field (F)	Origin	Food species
<i>Diplopsalis lenticula</i> Bergh	M2	U	Gulf of Mexico, FL	Db, Ca
<i>Preperidinium meunerii</i> (Pavillard) Elbrächter	n/a	F	Salt Pond, MA	n/a
<i>Protoperidinium conicoides</i> (Paulsen) Balech	n/a	F	Salt Pond, MA	n/a
<i>Protoperidinium crassipes</i> (Kofoid) Balech	M065-PC-C17	C	Gulf of Mexico, FL	Lp
<i>Protoperidinium depressum</i> (Bailey) Balech	PD1R1A-C11	C	SW Coast of Ireland	Db, Ca
<i>Protoperidinium oblongum</i> (Aurivillius) Parke and Dodge	24C(5)-PO-1	C	SW Coast of Ireland	Db, Ca
<i>Protoperidinium steidingerae</i> Balech	MV0923-PO-C17	C	Vineyard Sound, MA	Db, Ca
<i>Protoperidinium</i> sp. 1 Bergh	M064-sm-1	U	Gulf of Mexico, FL	n/a

sufficiently divergent paralogous genes are used in phylogenetic investigation, the result can be a tree that reflects the evolution of the gene sequence but not of the species (Sanderson & Doyle 1992; Buckler *et al.* 1997). For example, if different paralogs give rise to the major locus in two different species, the divergence indicated by the gene tree would be more ancestral than the speciation event. Additionally, gene families with intermediate levels of recombination or conversion between paralogs are unlikely to give the correct species phylogeny (Sanderson & Doyle 1992). Understanding the extent and type of intraspecific gene variability is thus critical for robust phylogenetic studies.

In this study, we characterized the intragenomic diversity in the D1–D6 region of the LSU rDNA gene in eight morphospecies of thecate, heterotrophic dinoflagellates, documenting extremely high variability within morphospecies and even within a single cell. We investigated the phylogeny of the polymorphic copies to determine whether the LSU rDNA gene remains an appropriate marker for phylogenetic investigation of these heterotrophic protists. Examination of LSU rDNA gene expression was conducted to determine whether polymorphic copies were functional genes or nonfunctional pseudogenes.

MATERIAL AND METHODS

Isolation and culturing

Six *Protoperidinium* species and two diplopsalid species were isolated from sites around the North Atlantic, including off the coast of Massachusetts, in the Gulf of Mexico, Florida, and from the southwestern coast of Ireland. Five of these *Protoperidinium* species and *Diplopsalis lenticula* Bergh were brought into culture, and the other two species were studied using specimens isolated from the field (Table 1).

Heterotrophic dinoflagellates were grown in 0.2- μ m-filtered, Teflon-autoclaved seawater from Vineyard Sound (30 psu, or amended to 35 psu by evaporation) and fed appropriate phytoplankton food species (Table 1). Cultures were maintained at 15°C under low light (*c.* 50 μ mol photon $m^{-2} s^{-1}$) on a 14 h:10 h light:dark cycle in 70 ml untreated tissue culture flasks (Falcon 353009, Becton Dickinson, Franklin Lakes, NJ) that were rotated on a plankton wheel at 1–2 rpm. Cultures were transferred every 4–5 days by pouring approximately two thirds the volume of the old culture into a new flask containing fresh sterile seawater and phytoplankton prey.

Autotrophic dinoflagellate cultures were maintained at 15°C or 20°C in tubes with 25 ml of f/2 nutrient medium minus silicate (Guillard 1975), except for *Lingulodinium polyedrum* (Stein) Dodge (GPES 22), which was grown in ES medium (Kokinis & Anderson 1995). Diatom cultures (CCMP 356 and CCMP 158) were maintained in tubes with 25 ml of f/2 nutrient medium plus silicate at 15°C. All prey cultures were kept at a photon flux density of *c.* 100 μ mol $m^{-2} s^{-1}$ on a 14 h:10 h light:dark cycle.

Species identification

Morphological species identification of cultured *Protoperidinium* species and *D. lenticula* was confirmed by analysis of thecal plate structure using scanning electron microscopy (SEM) or by epifluorescence microscopy of cells stained with Calcofluor White MR2 (Polysciences, Warrington, PA, USA). Species not brought into culture were identified by light microscopy. Individual live cells were isolated by micropipette from field samples; examined for shape, size, and thecal plate morphology on a Zeiss Axiovert S100 light microscope at $\times 100$ –200, and photographed using a Sony Exwave HAD 3CCD color video camera integrated with Scion Image 1.62 software. The remainder of the field sample was preserved with borate-buffered formalin (5% final concentration), and species identification of the live cells was verified by Calcofluor White staining of cells of similar morphology from the preserved field sample.

SEM: Samples were preserved with borate-buffered formalin (5% final concentration) and stored at 4°C at least overnight. Subsamples were centrifuged, aspirated to 1 ml, and brought up to 4 ml with filtered seawater. Several hundred heterotrophic dinoflagellate cells were separated from phytoplankton prey in the sample by micropipette isolation, deposited into 2-ml cryovials with 5% formalin in filtered seawater, and stored at 4°C overnight. Samples were drawn down onto filters (Nucleopore track-etched membrane, 13 mm, 5- μ m pore size) and washed first with filtered seawater and then with distilled, deionized water to remove fixatives and salts. Samples were dehydrated in a series of ethanol washes of increasing concentration, critical point dried (Tousimis Samdri-780A, Tousimis Research Corp., Rockville, MD, USA), sputter coated with gold palladium (Tousimis Samsputter-28, Tousimis Research), and examined on an SEM (JEOL JSM-840, Peabody, MA, USA).

CALCOFLUOR WHITE: Cultures and field samples were preserved with borate-buffered formalin (5% final concentration)

and stored at 4°C until analysis. Thecal plate staining was done with Calcofluor White according to the method of Fritz and Triemer (1985). Subsamples of stained cells were examined at $\times 100$ – 200 on a Zeiss Axioskop microscope with a 100-W mercury lamp and a Zeiss #2 filter set (excitation 365 nm, emission 420 nm). Images were taken with a Zeiss MC 100 digital camera system.

Single-cell PCR amplification, cloning, and sequencing of LSU rDNA

Single-cell PCR was used to amplify the target sequence as described in Gribble and Anderson (2006). Single *Protoperidinium* sp. cells were isolated from culture or field samples by micropipette, washed two to three times in sterile filtered seawater and one to two times in sterile DI water, and deposited individually into PCR tubes in 10 μ l of sterile DI water. To enhance cell lysis, PCR tubes with isolated cells were frozen at -80°C overnight and just before PCR were subjected to a sonification ice bath at 40 A for approximately 30 s.

The single cells were used directly as template to amplify approximately 1430 bp of the LSU rDNA containing the variable domains D1–D6, using the primers D1R (Scholin *et al.* 1994) and 28–1483R (Daugbjerg *et al.* 2000). The 50- μ l PCR reaction mixture contained 2.5 units of *Pfu*, a proofreading DNA polymerase (Stratagene, La Jolla, CA, USA), 5 μ l 10 \times buffer (1 \times final concentration), 0.3 μ M of each primer, and 200 μ M dNTPs (Takara, Shiga, Japan). Thermal cycling was conducted using an initial denaturation at 95°C for 5 min, 30 cycles of 95°C for 1 min, 50°C for 1 min, and 72°C for 2 min, followed by a final elongation step of 72°C for 10 min.

Between 25 and 30 μ l of PCR product were run on a 1% agarose gel. Positive bands were excised, and the product was purified and concentrated using a MinElute Gel Extraction Kit (Qiagen, Valencia, CA, USA). For each species, between three and six purified PCR products, each from a single cell, were cloned separately using the Zero Blunt TOPO PCR Cloning Kit for Sequencing (Invitrogen, Carlsbad, CA, USA). Primers T3, T7 and a species-specific internal primer (Gribble & Anderson 2006) were used for sequencing between 12 and 89 clones for each species. Sequencing was done on an Applied Biosystems 3730XL capillary sequencer. Sequences were deposited in GenBank under the accession numbers EF152778–EF152991.

rDNA gene expression

To determine if all LSU rDNA sequences for a given species were functional or if some were unexpressed pseudogenes, gene expression was examined in *Protoperidinium oblongum* (Aurivillius) Parke and Dodge, *Protoperidinium steidingeriae* Balech, and *Protoperidinium crassipes* (Kofoid) Balech using reverse-transcription PCR (RT-PCR). These three species were chosen, as they had been shown in this study to have relatively high levels of intraspecific diversity as measured by number of unique sequences, number of polymorphic sites, and/or number of indels.

Total RNA was extracted from between 1 and 50 *Protoperidinium* sp. cells using the Ambion RnAqueous MicroElute Kit (Ambion, Austin, TX, USA). Immediately after elution of RNA, samples were treated twice with DNase (37°C, 15 min) provided with the kit to remove any contaminating DNA, fol-

lowed by two treatments with DNase inactivation reagent. Reverse transcription of RNA to cDNA was according to the protocol accompanying SuperScript III (Invitrogen) using specific primers 28–1483R (Daugbjerg *et al.* 2000) or D3B (Nunn *et al.* 1996). Control reactions with no reverse transcriptase were always run to check for DNA contamination. After reverse transcription, samples were treated with RNase H (Invitrogen) to remove any RNA that might act as template in the subsequent PCR reactions. PCR reactions were carried out using the same reagent concentrations and cycle conditions as described above, except with an annealing temperature of 55°C instead of 50°C. Positive bands were excised, purified, cloned, and sequenced as described above.

Test for PCR error

In order to determine if PCR error was the cause of the observed intragenomic sequence diversity, four clones of PCR products from *P. steidingeriae* were reamplified by PCR, using the same D1R and 28–1483R primers employed for PCR throughout the study. The new PCR products were purified and sequenced directly. Additionally, all four purified new PCR products were cloned as described above, and 8 to 16 clones were sequenced for each PCR product. Since the initial clones were known to have a single version of the LSU rDNA sequence, any differences between the newly amplified, cloned, and sequenced PCR products could be attributed to PCR error due to misincorporation of an incorrect nucleotide or indel creation by *Pfu* polymerase or to replication error in cloning. Although this method did not allow differentiation between PCR error and cloning error, the ability to sequence purified PCR products directly, without cloning, should indicate a lack of high levels of PCR error.

Phylogenetic analysis

Partial LSU rDNA sequences from *Protoperidinium* spp., *D. lenticula*, and *Preperidinium meunerii* (Pavillard) Elbrächter were edited using Sequencher 4.5 (Gene Codes Corporation, Ann Arbor, MI, USA). Uncertain base calls were corrected by hand with reference to the sequence chromatograph. Edited sequences were aligned using Clustal X v.1.83.1 (Thompson *et al.* 1997), and resulting alignments were edited by hand in MacClade 4.07 (Sinauer Associates, Sunderland, MA, USA). Within a species, all sequences could be fully aligned. When sequences from multiple species were aligned, regions that could not be reliably aligned were excluded from analyses.

Pairwise differences were calculated using neighbor joining analysis (gaps treated as missing data) in PAUP 4.0b10 (Swofford 2002). Intraspecific pairwise differences were determined from alignments of all the LSU rDNA sequences for a given species. Pairwise differences between species were calculated from an alignment of the most abundant sequence in the clone library for each species, excluding sites 554–662, which were too divergent to be aligned.

The numbers and sizes of insertions/deletions (indels) were determined by manual examination of alignments in MacClade 4.07. In those cases where multiple indels of different sizes overlapped in the same region, they were counted as a single indel since they may have resulted from a single initial insertion or deletion that was subsequently expanded or reduced in other gene copies during replication. The number of

Table 2. Intraspecific variability in the D1–D6 region of LSU rDNA for eight species of thecate heterotrophic dinoflagellates. The number of polymorphic sites and the percent pairwise sequence difference do not include indels.

Species	No. individual cells	No. clones	No. unique sequences	No. polymorphic sites (%)	No. indels (size range bp)	Mean pairwise difference (%)	Standard deviation (%)	Maximum pairwise difference (%)	Minimum pairwise difference (%)
<i>Diplopsalis lenticula</i>	5	12	5	48 (3.5)	1 (1)	0.12	0.06	0.21	0.07
<i>Preperidinium meunerii</i>	3	32	11	15 (1.0)	0	0.21	0.11	0.49	0.07
<i>Protopteridinium conicoides</i>	3	89	13	17 (1.2)	0	0.18	0.09	0.42	0.07
<i>Protopteridinium crassipes</i>	5	43	32	78 (5.4)	9 (1–20)	0.98	0.55	2.0	0.07
<i>Protopteridinium</i> sp. 1	6	47	14	289 (20)	30 (1–95)	3.6	5.9	21.0	1.1
<i>Protopteridinium oblongum</i>	3	63	30	121 (8.4)	7 (1)	0.66	0.71	3.6	0.70
<i>Protopteridinium steidingerae</i>	3	89	33	21 (1.5)	2 (1–179)	0.24	0.15	0.64	0.08
<i>Protopteridinium depressum</i>	7	83	38	328 (22)	24 (1–63)	2.1	2.0	11.0	0.08

indels reported may therefore be an underestimate of the number of independent insertion or deletion events. The number of polymorphic sites across an alignment of intraspecific LSU rDNA sequences was determined using DnaSP 4.10.7 software (Rozas *et al.* 2003).

Neighbor joining analysis was used to create a phylogenetic tree of all 179 sequences from the eight heterotrophic dinoflagellates studied plus rDNA sequences of *Amphidinium carterae* (GenBank AY455669) and *Amphidinium steinii* (GenBank AY455673) as outgroups using PAUP* 4.0b10 (Swofford 2002). Identical sequences within a species were represented by a single sequence in the phylogenetic analysis. Nucleotide positions 554–662 were too divergent to be aligned and were excluded from the phylogenetic analysis, leaving a total of 1455 bases. Gaps were treated as missing data. Branch swapping was by tree-bisection-reconnection. Bootstrap support values were determined from 1000 replicates.

Rarefaction analysis

Rarefaction analysis was conducted to estimate whether the full diversity of LSU rDNA sequences had been sampled for each species. Rarefaction allows the calculation of sequence diversity for a given number of sampled clones. The number of clones sequenced for each sequence variant were input into Analytic Rarefaction 1.3 (Holland 2003) to produce rarefaction curves showing the expected number of unique sequences that would be found with increasing levels of sampling effort for a given species.

RESULTS

Sequencing directly from PCR products, using either purified or unpurified products, resulted in unresolved chromatograms, often with multiple peaks. When the PCR products were cloned and those clones were sequenced, multiple, diverse copies of the LSU rDNA were found, even within a single cell. The sequence clone library for each species was usually comprised of one highly represented sequence and many unique sequences (Table 2). Within any given species, differences between sequences were characterized primarily by single base pair substitutions and indels ranging in size from 1 bp up to 179 bp (e.g. Fig. 1). In a single sequence from *Pro-*

topteridinium depressum (Bailey) Balech, a microsatellite was inserted, with the bases TG repeated in tandem 10 times, interrupted by bases GG, and then repeated an additional 15 times. In all species, nucleotide substitutions and indels were scattered across the D1–D6 region, with no apparent clustering in either more or less divergent domains.

Cloned PCR products that were reamplified by PCR could easily be sequenced directly, suggesting that the inability to directly sequence PCR products from heterotrophic dinoflagellates was not due to PCR error. As additional evidence that the observed LSU rDNA sequence polymorphism was not due to PCR error, when four cloned PCR products were reamplified by PCR, the new PCR products cloned, and 8 to 16 of the clones from each reaction were sequenced, the sequences were 100% identical and were identical to those of directly sequenced PCR products.

The degree of intraspecific sequence diversity, quantified by pairwise differences between sequences and by the number of polymorphic sites in the alignment of sequences for a given species, varied among species (Table 2). Indels were not included in pairwise difference or polymorphic site analyses of intraspecific pairwise differences.

Protopteridinium conicoides Paulsen (Balech) and the lenticular dinoflagellates *D. lenticula* and *P. meunerii* had the lowest sequence diversities of the morphospecies studied, with mean intraspecific pairwise differences of 0.18%, 0.12%, and 0.21%, respectively. These species had relatively low percentages of polymorphic sites, with 1.0% polymorphic sites for *P. meunerii* and 1.2% for *P. conicoides*, although the percent polymorphic sites was slightly higher for *D. lenticula*, at 3.5%. *Preperidinium meunerii* and *P. conicoides* sequences had no indels, and *D. lenticula* had only a single indel of 1 bp.

Protopteridinium sp. 1 had the highest mean intraspecific pairwise difference (3.6%), with a maximum pairwise difference between any two sequences of 21.0%. *Protopteridinium depressum* and *P. oblongum* had relatively high mean pairwise differences and high numbers of polymorphic sites. Sequences from these species were characterized by many large indels. *Protopteridinium steidingerae* had relatively lower mean pairwise differences (0.24%) and numbers of polymorphic sites (1.5%) but contained a large indel of 179 bp. *Protopteridinium crassipes* had intermediate intraspecific sequence variability, with a mean pairwise difference of 0.98%, 5.4% polymorphic sites, and nine indels up to 20 bp in length.

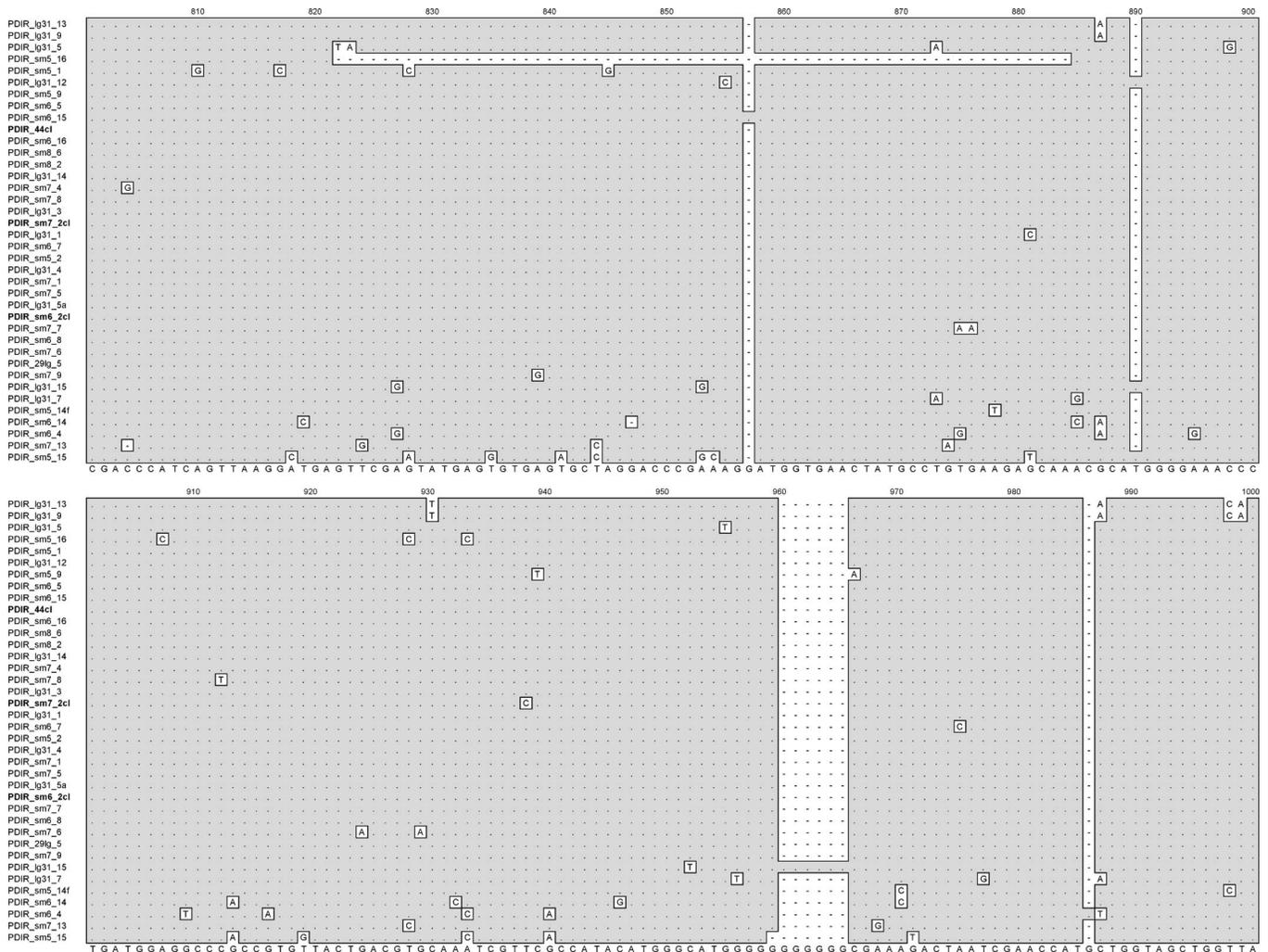


Fig. 1. Alignment of partial LSU rDNA sequences from a single clonal culture of *Protoperidinium depressum* (800–1000 bp). The library of 83 clones was created from seven separate single-cell PCR reactions. Sequences that represent two or more identical sequences are listed in bold. For example, sequence PDIR 44cl represents 44 identical clones from all seven PCR reactions.

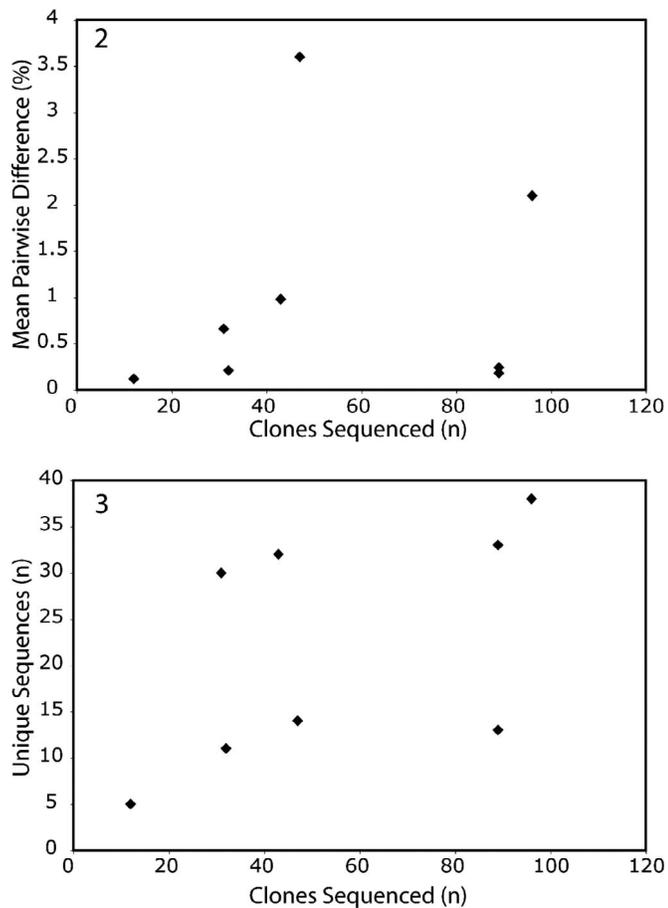
Rarefaction analysis indicated that we had not sampled the full diversity of LSU rDNA sequences for any species despite sequencing between 12 and 89 PCR product clones for each species (data not shown). For all morphospecies tested, rarefaction curves did not reach a plateau. This analysis does not allow determination of the total number of sequence variants for a given species, however. Additional sampling would have to be done to fully characterize all the different sequence variants within a given species.

Intraspecific sequence diversity, measured both by the number of unique sequences in the clone library for a given species and by the mean intraspecific pairwise difference, was not dependent upon the number of clones sequenced for a given species (Figs 2, 3). Thus, variability in the levels of sequence diversity among morphospecies seems to be inherent and is not due to differences in sampling effort for different species.

The mean intraspecific LSU rDNA sequence variability was in all cases lower than the mean interspecific variability (Table 3). The maximum pairwise differences of *Protoperidinium* sp. 1 (21.0%) and *P. depressum* (11.0%) were approaching the

differences between closely related species like the diplopsalids *D. lenticula* and *P. meunerii* and the Orthoperidinium species *P. depressum* and *P. oblongum*.

Despite the high intraspecific sequence variability, sequence variants for each morphospecies were monophyletic, a finding supported by high bootstrap values (Fig. 4). The level of sequence variability can be estimated by the lengths of the branches in the phylogenetic tree. Those morphospecies with similar levels of sequence diversity tend to be in the same clade. The diplopsalid species, which had very low intragenomic sequence diversity, were basal to the *Protoperidinium* species. The *Protoperidinium* split into two main clades. The first clade contained species of intermediate sequence diversity, including *P. conicoides*, *Protoperidinium* sp. 1, and *P. crassipes*, with *P. conicoides* branching as a sister group to *Protoperidinium* sp. 1 and *P. crassipes*. The second clade contained species of highest intraspecific diversity: *P. depressum*, *P. oblongum*, and *P. steidingerae*, with *P. steidingerae* branching separately from the more closely related *P. depressum* and *P. oblongum*.



Figs 2, 3. For each of the eight species analyzed, comparison of number of clones sequenced with (Fig. 2) mean pairwise difference (%), and (Fig. 3) number of unique sequences. No relationship between cloning effort and observed intraspecific diversity was found ($r^2 < 0.3$ for regression analyses in both cases).

LSU rDNA gene expression

The diversity of expressed LSU rDNA sequences was lower than the total diversity of genomic sequences in each of the three species analyzed by RT-PCR (Table 4). *Protoperidinium oblongum*, for example, which had a mean pairwise difference

of 0.66% among polymorphic sequences amplified from genomic DNA, exhibited a mean pairwise difference in expressed sequences of 0.14%, with a maximum pairwise difference of 0.21%. While 30 of 31 sequence clones (97%) from genomic DNA were unique, only 4 of 17 clones (23%) of expressed sequences were unique in *P. oblongum*. Additionally, there were no indels in expressed sequences of *P. oblongum*, compared with seven single base pair indels in clones from genomic DNA.

Protoperidinium crassipes and *P. steidingerae* expressed a larger portion of their complement of LSU rDNA genes than did *P. oblongum*. In *P. crassipes*, 15 of 37 clones (41%) of expressed sequences were unique, down from the 32 unique clones out of 43 total (74%) in genomic clone libraries. Expressed sequences of *P. crassipes* had fewer polymorphic sites and no indels, compared with sequences from genomic DNA. *Protoperidinium steidingerae* expressed 16 unique sequences from 31 clones (51%), up slightly from 33 unique out of 89 total clones from genomic DNA (39%). The expressed sequences in *P. steidingerae* had approximately the same percent of polymorphic sites and smaller indel sizes than sequences from the genomic clone library.

Sequences found in RT-PCR overlapped with those from the clone libraries from genomic DNA for all three species. For *P. crassipes*, 13% of sequences from the genomic clone library overlapped with 27% of the sequences from RT-PCR. In *P. oblongum*, 3% of genomic sequences were identical to 25% of RT-PCR sequences, and in *P. steidingerae*, 9% of genomic sequences matched 18% of sequences from RT-PCR.

Results of analytic rarefaction analysis suggested that not all expressed copies of the LSU rDNA were sequenced for any of the three species. For *P. crassipes*, *P. oblongum*, and *P. steidingerae*, the rarefaction curves did not reach a maximum horizontal asymptote (data not shown).

DISCUSSION

Thecate heterotrophic dinoflagellate species investigated here were found to have higher-than-expected intraindividual, intraspecific, and interspecific diversity in LSU rDNA sequences. The level of sequence diversity varied among morphospe-

Table 3. Interspecific variability, shown as percent pairwise difference not including indels, in the D1–D6 region of LSU rDNA between eight species of thecate heterotrophic dinoflagellates.

	<i>Preperidinium meuneri</i>	<i>Diplopsalis lenticula</i>	<i>Protoperidinium crassipes</i>	<i>Protoperidinium sp. 1</i>	<i>Protoperidinium conicoides</i>	<i>Protoperidinium steidingerae</i>	<i>Protoperidinium depressum</i>
<i>Diplopsalis lenticula</i>	19.0						
<i>Protoperidinium crassipes</i>	30.8	34.3					
<i>Protoperidinium sp. 1</i>	25.0	26.6	26.6				
<i>Protoperidinium conicoides</i>	30.4	32.4	35.6	29.4			
<i>Protoperidinium steidingerae</i>	26.8	30.1	36.7	33.3	36.3		
<i>Protoperidinium depressum</i>	29.1	29.9	39.0	33.0	40.2	34.8	
<i>Protoperidinium oblongum</i>	30.0	31.7	40.0	34.2	39.4	35.0	8.2

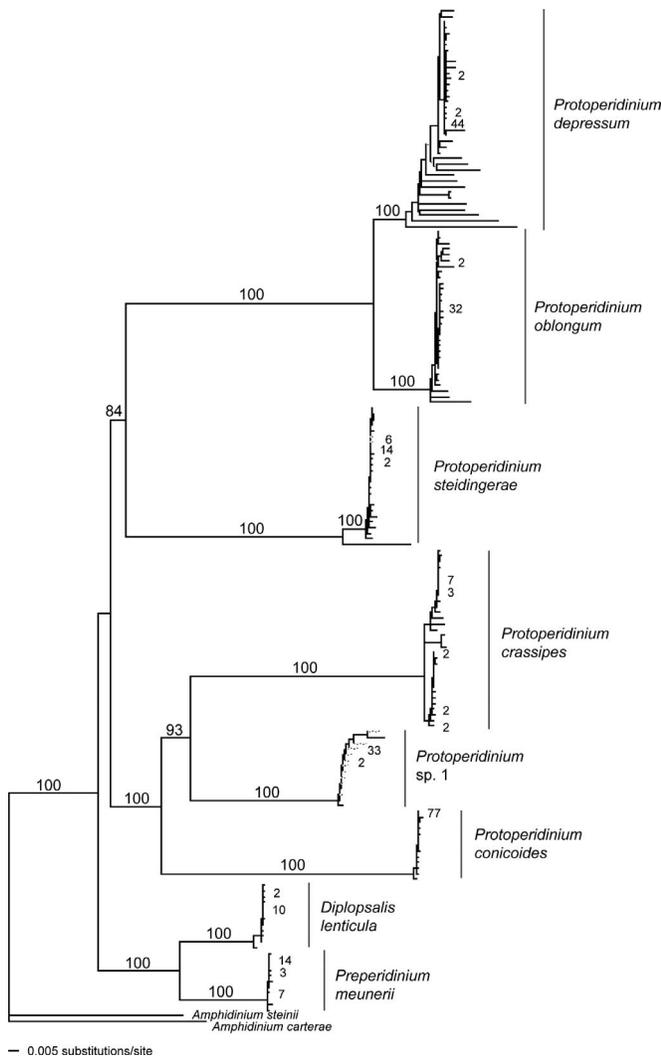


Fig. 4. Neighbor joining tree of heterotrophic dinoflagellate phylogeny inferred from the D1–D6 region of LSU rDNA with *Amphidinium steinii* Lemmermann (GenBank accession AY455673) and *Amphidinium carterae* Hulbert (GenBank accession AY455669) as outgroups. Sequences used for each species were from clone libraries created from three to seven separate single-cell PCR reactions. Gaps were treated as missing data. Neighbor joining bootstrap support values from 1000 replicates are shown above the node. Bootstrap support values lower than 50% are not shown. Numbers at the ends of terminal nodes refer to the number of clones sharing the sequence represented by that node.

cies. Some, but not all, of the different copies are nonfunctional pseudogenes. These species appear to express more than one copy of the LSU rDNA, although the expressed sequences were generally less variable than the genomic copies.

Results of previous studies have indirectly indicated that *Proto-peridinium* species might have high intraspecific rDNA sequence polymorphism. Reports of difficulties in direct sequencing of PCR products of LSU rDNA genes from *Proto-peridinium* spp. (e.g. Bolch 2001) were likely due to the high degree of intraspecific and intraindividual sequence diversity documented here. In an investigation of the phylogeny of the *Proto-peridinium* inferred from the SSU rDNA, where sequencing was done directly from PCR products, a high sequence variability among conspecific individuals taken from the same field sample was reported (Yamaguchi & Horiguchi 2005). In that study, because specimens were taken from field samples rather than clonally cultured, and PCR products were not cloned, a distinction could not be made between intraspecific diversity (cryptic species) and possible intraindividual sequence diversity. In SSU and LSU rDNA phylogenies for dinoflagellates that included *Proto-peridinium* and diplopsalids (Saldarriaga *et al.* 2004; Gribble & Anderson 2006), *Proto-peridinium* was one of the most divergent genera compared with most other dinoflagellates, suggesting a higher rate of evolution of the rDNA in this lineage of heterotrophic dinoflagellates. A higher rate of evolution could be the cause of the high intraspecific and intraindividual rDNA sequence variability found in this study.

Intraspecific heterogeneity in the ribosomal genes of several dinoflagellate species has been reported previously (e.g. Rowan & Powers 1992; Scholin *et al.* 1993; Yeung *et al.* 1996; Rehnstam-Holm *et al.* 2002; Montresor *et al.* 2003; Santos *et al.* 2003; Kim *et al.* 2004). In most of those cases, researchers found only one or two polymorphic copies, a much lower degree of intraspecific polymorphism than that discovered here. In other cases, particularly in studies of the ITS region, several sequence variants were found within a clonal dinoflagellate culture (Rowan & Powers 1992; Montresor *et al.* 2003). In the earlier studies, fewer sequenced clones were analyzed, possibly accounting for some of the difference between our study and the previous research. Large sample sizes may be needed for accurate estimates of sequence diversity in a species.

The level of intraindividual sequence polymorphism seen here initially raised concern about the role of PCR error in artificially creating sequence diversity. Recombination of sequences in multigene families like rDNA genes can occur during PCR, resulting in hybrid sequences and causing an overestimate of sequence diversity. Such recombination combines already existing sequences, however, and does not generate

Table 4. Intraspecific variability in the D1–D6 region of expressed LSU rDNA sequences for three species of *Proto-peridinium*, as determined by RT-PCR. The number of polymorphic sites and the percent pairwise difference do not include indels.

Species	No. single-cell PCR reactions	No. clones	Unique sequences (n)	No. polymorphic sites (%)	No. indels (size range bp)	Mean pairwise difference (%)	Standard deviation (%)	Maximum pairwise difference (%)	Minimum pairwise difference (%)
<i>Proto-peridinium crassipes</i>	3	37	15	18 (1.9)	0	0.30	0.13	0.64	0.11
<i>Proto-peridinium oblongum</i>	2	17	4	4 (0.28)	0	0.14	0.06	0.21	0.07
<i>Proto-peridinium steidingerae</i>	4	31	16	12 (1.2)	4 (1–14)	0.18	0.10	0.50	0.1

polymorphisms *de novo*. Although chimeric sequences were not apparent in the data, we could not rule out some level of PCR-induced recombination.

Error can also occur during PCR because of incorporation of the wrong nucleotide by DNA polymerase during amplification. To minimize enzyme-mediated PCR error, we switched from *Taq* DNA polymerase, with an error rate estimated at 8.0×10^{-6} errors per nt (Cline *et al.* 1996) to 1.1×10^{-4} errors per nt (Cariello *et al.* 1991; Barnes 1992), to the proof-reading DNA polymerase, *Pfu*, with an error rate of 1.3×10^{-6} errors per nt (Cline *et al.* 1996). This change resulted in no appreciable difference in the level of intraindividual or intraspecific LSU rDNA polymorphism, indicating that the polymorphism was likely real and not due to enzyme replication error during PCR. Additionally, there was no difference in the degree of sequence polymorphism between PCR reactions conducted with DNA extracted from individual or multiple cells or from those done directly on isolated single cells without DNA extraction (data not shown), indicating that PCR error due to our single-cell PCR method was not the cause of the high degree of intraindividual or intraspecific sequence diversity. A check for PCR error by reamplifying clones of known sequence and both direct sequencing and cloning and then sequencing the clones resulted in 100% identity between all sequences, suggesting that PCR error or replication error during cloning was negligible.

The match between sequences found in the genomic clone libraries and the RT-PCR clone libraries, ranging from 3% to 33% identity, provides another indication that most of the sequence variability observed was not due to PCR error. The same PCR errors would be unlikely to occur during separate PCR reactions run under different conditions.

PCR error cannot be discounted as a source of some variability, however. In all species, the lowest pairwise differences may be accounted for by PCR error, as they fall within the error range for *Pfu* polymerase. In the case of the species with lowest rates of difference, *D. lenticula*, *P. meunerii*, and *P. conicoides*, PCR error could account for the majority of polymorphic sequences.

The current work indicates that there are significant differences among species in the level of intraindividual and intraspecific variability in the LSU rDNA. Mean pairwise difference was 30 times greater in *P. crassipes* than in *P. meunerii*, for example. Gene duplication and concerted evolution appear to be acting at different rates in various clades of dinoflagellates. Intraindividual sequence variability in ribosomal genes has been found to different degrees in organisms ranging from salmonid fish (Reed *et al.* 2000) to humans (Brownell *et al.* 1983) to other dinoflagellate species (Montresor *et al.* 2003; Kim *et al.* 2004) and could be due to high mutation rates or to ancestral hybridization events between closely related species or strains. Maintenance of sequence diversity could allow faster adaptation to environmental conditions and provide template for accelerated evolution. A better understanding of the evolutionary mechanisms responsible for the creation and maintenance of such sequence diversity could expand the tools available for phylogenetic investigation.

Intragenomic sequence variability will accumulate if the rate of gene conversion is lower than the rate of mutation or if concerted evolution is slower than speciation. High intraspecific sequence variability may also exist if a species is in

a 'transitional state' of concerted evolution (Elder & Turner 1995). In the heterotrophic dinoflagellates examined here, greater sequence similarity of the LSU rDNA within a species than between species implies a mechanism to maintain sequence homogeneity. The high degree of intraspecific sequence polymorphism indicates that concerted evolution is acting more slowly than mutations are occurring in the *Protoperidinium* spp. examined, however. Nothing is known about the chromosomal location of the ribosomal genes in dinoflagellates, but given the large DNA content ($3\text{--}250$ pg cell⁻¹; Spector 1984) and many chromosomes in dinoflagellates (e.g. 143 chromosomes in *Alexandrium fundyense*, reviewed in Hackett *et al.* 2004), it is likely that they bear high numbers of rDNA loci on multiple and possibly nonhomologous chromosomes, limiting the effectiveness of concerted evolution. Intrachromosomal homogenization of repeated sequences is faster than interchromosomal recombination (Liao *et al.* 1997; Polanco *et al.* 1998). Additionally, effective homogenization of dispersed genes by gene conversion, that is, the nonreciprocal exchange of genetic information between similar sequences, is only effective when the number of dispersed genes is small (reviewed by Liao 1999). Concerted evolution does not appear to be slower than speciation, however, since divergent paralogs were not found. Differences between mutation rates and DNA repair rates can also lead to an accumulation of sequence variants in genes with multiple copies (Elder & Turner 1995). Our results do not allow us to distinguish between the possible causes of the observed intraspecific variability.

Interspecific LSU rDNA sequence diversity was greater than intraspecific diversity in all species studied and in phylogenetic analysis the sequences for each species group together as a single clade of apparently paralogous genes. The phylogeny created from all sequenced clones has the same species topology as does the phylogeny created using a single sequence for each species (Gribble & Anderson 2006). Care must be used in employing ribosomal genes as phylogenetic markers, however, as the ability to distinguish between paralogs and orthologs becomes difficult with increasing sequence polymorphism. As more closely related species are sequenced or attempts are made to investigate cryptic species, divergent paralogs might be revealed, and use of the LSU rDNA in phylogenetic reconstruction may become more complicated.

Scattering of changes through both variable and conserved regions of the LSU rDNA in the heterotrophic dinoflagellates examined initially seemed to indicate that most sequences might not be functional. RT-PCR confirmed that the diversity of sequences in a given species was comprised of both pseudogenes and expressed sequences. The expressed sequences in *P. crassipes*, *P. oblongum*, and *P. steidingeriae* were less diverse than were the pseudogenes in these species and generally had fewer and smaller indels. Because rRNAs are structural molecules that must be abundant for protein translation and cell function, it could be deleterious to simultaneously express heterogeneous rRNAs. On the other hand, expression of more than one version of the LSU rDNA could be beneficial, perhaps allowing adaptability to changing environmental conditions.

CONCLUSIONS

The results of this study have important implications for environmental sequencing studies in which DNA extracted directly from environmental samples is amplified by PCR and used to create clone libraries. Such projects may greatly overestimate the species diversity in samples if unique phylotypes are assumed to be individual species. In this study, we found pairwise sequence differences of up to 21.0%, not including differences due to indels, within a single culture. This level of sequence variability could easily be characterized as having been derived from separate species in environmental sequencing studies, in which the sequence difference for defining separate species is often set at about 3% (e.g. Venter *et al.* 2004).

These results indicate that polymorphism of ribosomal genes may be greater and more widespread among the dinoflagellates than is currently realized and provide evidence that the degree of intragenomic variability is not uniform across all dinoflagellates. Cloning and sequencing higher numbers of clones of PCR products from additional strains of the heterotrophic species studied here and from other autotrophic dinoflagellates would establish whether there is a significant difference in the level of intraspecific sequence polymorphism among various phylogenetic lineages of dinoflagellates. Finally, although ribosomal genes have long been standard tools in molecular phylogenetic investigations, our findings illustrate the need for thorough investigation of the genes chosen for inference of species phylogeny, as the evolution of the chosen gene may not perfectly mirror the evolution of the species of interest.

ACKNOWLEDGEMENTS

We thank the Marine Institute of Ireland, Trish Blum of the Mote Marine Laboratory, Florida, and David Kulis and Kenneth Houtler of the Woods Hole Oceanographic Institution for sample collection. Thank you to Dr. Malte Elbrächter for assistance with identification of *P. steidingerae*. Thank you to Dr. Mark Hahn for valuable comments. Funding for this project came from the Comer Foundation, the Cove Point Foundation, and National Science Foundation grant OCE-0136861.

REFERENCES

- AGUILERA A. & KLEIN H.L. 1989. Yeast interchromosomal recombination: long gene conversion tracts are preferentially associated with reciprocal exchange and require the *RAD1* and *RAD3* gene products. *Genetics* 123: 683–694.
- ALVERSON A.J. & KOLNICK L. 2005. Intragenomic nucleotide polymorphism among small subunit (18S) rDNA paralogs in the diatom genus *Skeletonema* (Bacillariophyta). *Journal of Phycology* 41: 1248–1257.
- ARNHEIM N., KRYSTAL M., SCHMICKEL R., WILSON G., RYDER O. & ZIMMER E. 1980. Molecular evidence for exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. *Proceedings of the National Academy of Sciences of the United States of America* 77: 7323–7327.
- BALECH E. 1974. El genero *Protoperidinium* Bergh, 1881 (*Peridinium* Ehrenberg, 1831, Partim). *Revista del Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" e Instituto Nacional de Investigacion de las Ciencias Naturales* 4: 1–79.
- BARNES W.M. 1992. The fidelity of Taq polymerase catalyzing PCR is improved by an N-terminal deletion. *Gene* 112: 29–35.
- BOLCH C.J.S. 2001. PCR protocols for genetic identification of dinoflagellates directly from single cysts and plankton cells. *Phycologia* 40: 162–167.
- BROWNELL E., KRYSTAL M. & ARNHEIM N. 1983. Structure and evolution of human and African ape rDNA pseudogenes. *Molecular Biology and Evolution* 1: 29–37.
- BUCKLER E.S. IV, IPPOLITO A. & HOLTSFORD T.P. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* 145: 821–832.
- CAMPBELL C.S., WOJCIECHOWSKI M.F., BALDWIN B.G., ALICE L.A. & DONOGHUE M.J. 1997. Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). *Molecular Biology and Evolution* 14: 81–90.
- CARIELLO N.F., SWENBERG J.A. & SKOPEK T.R. 1991. Fidelity of *Thermococcus litoralis* DNA polymerase (Vent) in PCR determined by denaturing gradient gel-electrophoresis. *Nucleic Acids Research* 19: 4193–4198.
- CLINE J., BRAMAN J.C. & HOGREFE H.H. 1996. PCR fidelity of *Pfu* DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Research* 24: 3546–3551.
- DAUGBJERG N., HANSEN G., LARSEN J. & MOESTRUP Ø. 2000. Phylogeny of some of the major genera of dinoflagellates based on ultrastructure and partial LSU rDNA sequence data, including the erection of three new genera of unarmoured dinoflagellates. *Phycologia* 39: 302–317.
- DOVER G. 1982. Molecular drive: a cohesive mode of species evolution. *Nature* 299: 111–116.
- ELDER J., JOHN F. & TURNER B.J. 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. *Quarterly Review of Biology* 70: 297–320.
- FRITZ L. & TRIEMER R.E. 1985. A rapid simple technique utilizing Calcofluor white M2R for the visualization of dinoflagellate thecal plates. *Journal of Phycology* 21: 662–664.
- GAINES G. & ELBRÄCHTER M. 1987. Heterotrophic nutrition. In: *The biology of dinoflagellates* (Ed. by F.J.R. Taylor), pp. 225–268. Blackwell Scientific Publications, Boston.
- GAINES G. & TAYLOR F.J.R. 1984. Extracellular digestion in marine dinoflagellates. *Journal of Plankton Research* 6: 1057–1062.
- GERNANDT D.S., LISTON A. & PIÑERO D. 2001. Variation in the nrDNA ITS of *Pinus* subsection *Cembroides*: implications for molecular systematic studies of pine species complexes. *Molecular Phylogenetics and Evolution* 21: 449–467.
- GRIBBLE K.E. & ANDERSON D.M. 2006. Molecular phylogeny of the heterotrophic dinoflagellates, *Protoperidinium*, *Diplopsalis*, and *Preperidinium* (Dinophyceae), inferred from LSU ribosomal DNA. *Journal of Phycology* 42: 1081–1095.
- GUILLARD R.R.L. 1975. Culture of phytoplankton for feeding marine invertebrates. In: *Culture of marine invertebrates* (Ed. by W.L. Smith & M.H. Chanley), pp. 29–40. Plenum Publishing, New York.
- HACKETT J.D., ANDERSON D.M., ERDNER D.L. & BHATTACHARYA D. 2004. Dinoflagellates: a remarkable evolutionary experiment. *American Journal of Botany* 91: 1523–1534.
- HARTMANN S., NASON J.D. & BHATTACHARYA D. 2001. Extensive ribosomal DNA genic variation in the columnar cactus *Lophocereus*. *Journal of Molecular Evolution* 53: 124–134.
- HILLIS D.M. & DIXON M.T. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarterly Review of Biology* 66: 411–453.
- HOLLAND S.M. 2003. Analytic Rarefaction 1.3. University of Georgia Stratigraphy Lab. Available at <http://www.uga.edu/~strata/software/> (27 May 2006).
- JACOBSON D.M. & ANDERSON D.M. 1986. Thecate heterotrophic dinoflagellates: feeding behavior and mechanisms. *Journal of Phycology* 22: 249–258.
- KIM E., WILCOX L., GRAHAM L. & GRAHAM J. 2004. Genetically distinct populations of the dinoflagellate *Peridinium limbatum* in neighboring northern Wisconsin lakes. *Microbial Ecology* 48: 521–527.

- KOKINOS J.P. & ANDERSON D.M. 1995. Morphological development of resting cysts in cultures of the marine dinoflagellate *Lingulodinium polyedrum* (= *L. machaerophorum*). *Palynology* 19: 143–166.
- LIAO D. 1999. Concerted evolution: molecular mechanism and biological implications. *American Journal of Human Genetics* 64: 24–30.
- LIAO D., PAVELITZ T., KIDD J., KIDD K. & WEINER A. 1997. Concerted evolution of the tandemly repeated genes encoding human U2 snRNA (the *RNU2* locus) involves rapid intrachromosomal homogenization and rare interchromosomal gene conversion. *EMBO Journal* 16: 588–598.
- MÁRQUEZ L.M., MILLER D.J., MACKENZIE J.B. & VAN OPPEN M.J.H. 2003. Pseudogenes contribute to the extreme diversity of nuclear ribosomal DNA in the hard coral *Acropora*. *Molecular Biology and Evolution* 20: 1077–1086.
- MONTRESOR M., SGROSSO S., PROCACCINI G. & KOOISTRA W. 2003. Intraspecific diversity in *Scrippsiella trochoidea* (Dinophyceae): evidence for cryptic species. *Phycologia* 42: 56–70.
- NAUSTVOLL L.J. 2000. Prey size spectra and food preferences in thecate heterotrophic dinoflagellates. *Phycologia* 39: 187–198.
- NUNN G.B., THEISEN B.F., CHRISTENSEN B. & ARCTANDER P. 1996. Simplicity-correlated size growth of the nuclear 28S ribosomal RNA D3 expansion segment in the crustacean order Isopoda. *Journal of Molecular Evolution* 42: 211–223.
- POLANCO C., GONZALEZ A., DE LA FUENTE A. & DOVER G. 1998. Multigene family of ribosomal DNA in *Drosophila melanogaster* reveals contrasting patterns of homogenization for IGS and ITS spacer regions: a possible mechanism to resolve this paradox. *Genetics* 149: 243–256.
- REED K.M., HACKETT J.D. & PHILLIPS R.B. 2000. Comparative analysis of intra-individual and inter-species DNA sequence variation in salmonid ribosomal DNA cistrons. *Gene* 249: 115–125.
- REHNSTAM-HOLM A.S., GODHE A. & ANDERSON D.M. 2002. Molecular studies of *Dinophysis* (Dinophyceae) species from Sweden and North America. *Phycologia* 41: 348–357.
- ROWAN R. & POWERS D.A. 1992. Ribosomal RNA sequences and the diversity of symbiotic dinoflagellates (zooxanthellae). *Proceedings of the National Academy of Sciences of the United States of America* 89: 3639–3643.
- ROZAS J., SÁNCHEZ-DELBARRIO J.C., MESSEGUER X. & ROZAS R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2497.
- SALDARRIAGA J.F., TAYLOR F.J.R., CAVALIER-SMITH T., MENDEN-DEUER S. & KEELING P.J. 2004. Molecular data and the evolutionary history of the dinoflagellates. *European Journal of Protistology* 40: 85–111.
- SANDERSON M.J. & DOYLE J.J. 1992. Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. *Systematic Biology* 41: 4–17.
- SANTOS S.R., KINZIE R.A., SAKAI K. & COFFROTH M.A. 2003. Molecular characterization of nuclear small subunit (18S)-rDNA pseudogenes in a symbiotic dinoflagellate (*Symbiodinium*, Dinophyta). *Journal of Eukaryotic Microbiology* 50: 417–421.
- SCHOLIN C.A., ANDERSON D.M. & SOGIN M.L. 1993. Two distinct small-subunit ribosomal RNA genes in the North American toxic dinoflagellate *Alexandrium fundyense* (Dinophyceae). *Journal of Phycology* 29: 209–216.
- SCHOLIN C.A., HERZOG M., SOGIN M. & ANDERSON D.M. 1994. Identification of group- and strain-specific genetic markers for globally distributed *Alexandrium* (Dinophyceae). II. Sequence analysis of a fragment of the LSU rRNA gene. *Journal of Phycology* 30: 999–1011.
- SPECTOR D.L. 1984. Dinoflagellate nuclei. In: *Dinoflagellates* (Ed. by D.L. Spector), pp. 107–147. Academic Press, Orlando, FL.
- SWOFFORD D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods), Version 4.0b10. Sinauer Associates, Sunderland, MA.
- THOMPSON J.D., GIBSON T.J., PLEWNIAC F., JEANMOUGIN F. & HIGGINS D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 22: 4673–4680.
- VENTER J.C., REMINGTON K., HEIDELBERG J.F., HALPERN A.L., RUSCH D., EISEN J., WU D., PAULSEN I., NELSON K.E., NELSON W., FOUTS D.E., LEVY S., KNAP A.H., LOMAS M.W., NEALSON K., WHITE O., PETERSON J., HOFFMAN J., PARSONS R., BADEN-TILLSON H., PFANNKÖCH C., ROGERS Y.-H. & SMITH H.O. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304: 66–74.
- YAMAGUCHI A. & HORIGUCHI T. 2005. Molecular phylogenetic study of the heterotrophic dinoflagellate genus *Protoperdinium* (Dinophyceae) inferred from small subunit rRNA gene sequences. *Phycological Research* 53: 30–42.
- YEUNG P.K.K., KONG K.F., WONG F.T.W. & WONG J.T.Y. 1996. Sequence data for two large-subunit rRNA genes from an Asian strain of *Alexandrium catenella*. *Applied and Environmental Microbiology* 62: 4199–4201.

Received 23 August 2006; accepted 19 January 2007
Associate editor: Niels Daugbjerg