J. Phycol. **43,** 344–355 (2007)
© 2007 No claim to original US government works
DOI: 10.1111/j.1529-8817.2007.00320.x

RECOGNIZING DINOFLAGELLATE SPECIES USING ITS rDNA SEQUENCES[1]

R. Wayne Litaker[2], *Mark W. Vandersea, Steven R. Kibler*

Center for Coastal Fisheries and Habitat Research, National Ocean Service, NOAA, 101 Pivers Island Road, Beaufort, North Carolina 28516-9722, USA

Kimberly S. Reece, Nancy A. Stokes

Virginia Institute of Marine Science, The College of William and Mary, P.O. Box 1346, Gloucester Point, Virginia 23062-1346, USA

François M. Lutzoni

Duke University, Department of Biology, Box 90338, Durham, North Carolina 27708, USA

Bryan A. Yonish

Duke University, Duke University Medical Center, Purple Zone, Durham, North Carolina 27708-3967, USA

Melissa A. West, Megan N. D. Black, and Patricia A. Tester

Center for Coastal Fisheries and Habitat Research, National Ocean Service, NOAA, 101 Pivers Island Road, Beaufort, North Carolina 28516-9722, USA

Dinoflagellate taxonomy is based primarily on morphology and morphometric data that can be difficult to obtain. In contrast, molecular data can be rapidly and cost-effectively acquired, which has led to a rapid accumulation of sequence data in GenBank. Currently there are no systematic criteria for utilizing taxonomically unassigned sequence data to identify putative species that could in turn serve as a basis for testable hypotheses concerning the taxonomy, diversity, distribution, and toxicity of these organisms. The goal of this research was to evaluate whether simple, uncorrected genetic distances (p) calculated using ITS1/5.8S/ITS2 (ITS region) rDNA sequences could be used to develop criteria for recognizing putative species before formal morphological evaluation and classification. The current analysis used sequences from 81 dinoflagellate species belonging to 14 genera. For this diverse assemblage of dinoflagellate species, the within-species genetic distances between ITS region copies ($p = 0.000–0.021$ substitutions per site) were consistently less than those observed between species ($p = 0.042–0.580$). Our results indicate that a between-species uncorrected genetic distance of $p \geq 0.04$ could be used to delineate most free-living dinoflagellate species. Recently evolved species, however, may have ITS p values <0.04 and would require more extensive morphological and genetic analyses to resolve. For most species, the sequence of the dominant ITS region allele has the potential to serve as a unique species-specific "DNA barcode" that could be used for the rapid identification of dinoflagellates in field and laboratory studies.

Key index words: **dinoflagellate; evolution; molecular species identification; morphospecies; ribosomal DNA**

Abbreviations: **bp, base pairs; ITS, internal transcribed spacer; ML, maximum likelihood; NTS, nontranscribed spacer; p, uncorrected genetic distance; PLOs, *Pfiesteria*-like organisms; QPCR, quantitative PCR; RGC, ribosomal gene complex**

Dinoflagellate species are described based on morphological differences, most commonly the number, shape, and size of the thecal plates covering the cell (Fensome et al. 1999). These morphological features are generally characterized using time-consuming electron microscopy. In contrast to morphological techniques, DNA sequencing methods have improved greatly in recent years, both in terms of accuracy and efficiency, making molecular characterization of cells technically easier and less expensive to perform (Haley et al. 1999, LaJeunesse and Trench 2000, Godhe et al. 2002, Galluzzi et al. 2004). These advances have resulted in the rapid accumulation of rDNA sequence data in GenBank. The frequency and diversity of unclassified rDNA dinoflagellate sequences in GenBank suggests the existence of numerous species that have yet to be morphologically evaluated. A method for systematically organizing these new sequence data into putative species groups could therefore help identify species before formal description. Further, the identification of putative species groups would assist in gen-

[1]Received 8 November 2005. Accepted 1 December 2006.
The U.S. government has the right to retain a nonexclusive royalty-free license in and to any copyright covering this paper.
[2]Author for correspondence: e-mail Wayne.Litaker@noaa.gov.

erating hypotheses concerning taxonomic relationships among dinoflagellates. Such an approach is particularly relevant, given that most extant dinoflagellates have yet to be described (Sournia 1986) and that the first potential identifier of new species will likely be an rDNA sequence.

This study specifically examined whether an empirical categorization system could be developed that would allow novel ITS1/5.8S/ITS2 (ITS region) rDNA sequence data to be partitioned into putative species groups. The method involved the following: (1) identifying ITS region sequence data from well-defined species, (2) calculating uncorrected genetic distance values (p) to estimate within- and between-species divergence, and (3) determining if a p value above a specified threshold correlated with recognized species. The ITS region was selected because there were sufficient data to undertake the study and because the region diverges rapidly during speciation. Consequently, this region has been successfully used to identify species as diverse as fungi (Larsson and Jacobsson 2004), vascular plants (Wagstaff 2004), and insects (Wagener et al. 2004). It should be emphasized that this approach is not intended to supplant traditional morphological species definitions, but rather to supplement morphological data and allow greater utilization of information on species diversity as new ITS region data are added to GenBank.

The first dinoflagellate ITS region was obtained from *Prorocentrum micans* Ehrenb. in the mid-1980s (Maroteaux et al. 1985). By the mid-1990s, ITS sequence data were being used to identify a number of harmful algal species (Adachi et al. 1996, 1997, Hudson and Adlard 1996). Since that time, species-specific molecular assays based on unique ITS sequences have been used to address important taxonomic, phylogenetic, and ecological questions concerning dinoflagellates (e.g., D'Onofrio et al. 1999, Penna and Magnani 1999, LaJeunesse and Trench 2000, Cho et al. 2001, Edvardsen et al. 2003, Galluzzi et al. 2004). Despite the wide use of ITS sequence data, no systematic study has been undertaken to determine if these sequences could be used to recognize putative species across a wide range of dinoflagellate groups.

Because ITS regions diverge rapidly, and vary in length, it is often difficult to align and compare homologous base pairs (bp). The accurate comparison of homologous bp is the most fundamental assumption underlying algorithms used to calculate phylogenetic relationships (Rosenberg 2005). Because of this alignment issue, we used uniform alignment parameters and then calculated uncorrected genetic distances as an alternative to phylogenetic analyses for discriminating dinoflagellate species. For this approach to be effective, the within-species variation must be systematically lower than the between-species variation, even when the alignments are suboptimal. A potential complication in assessing the within-species ITS region sequence variation is that there are multiple rDNA copies per genome. Estimates range from 200 to 1200 copies per cell, depending on the species (Maroteaux et al. 1985, Galluzzi et al. 2004). Random sequence mutations can occur in the rDNA copies, resulting in sequence variation both among ITS region copies in the same genome as well as between genomes of the same species (Le Blancq et al. 1997). In this study, we calculated both intragenomic (within isolate) and intergenomic (between individual isolates) genetic distances to determine if sequence variation was equivalent within and among genomes of the same species.

Some heterogeneity among ITS copies is expected since they serve as spacer regions within the ribosomal gene complex (Nazar 2004). As such, these regions would be more likely to accumulate mutations than the flanking small subunit (SSU), 5.8S, and large subunit (LSU) genes, which are under strong stabilizing selection because of their critical role in protein synthesis (Lygerou et al. 1996, Allmang et al. 1999, 2000). However, the frequency of these mutations within a species is reduced by concerted evolution, the process by which gene conversion and unequal crossover events maintain the integrity of multigene families (Ohta 2000, Kovarik et al. 2004, Rooney 2004).

A complex interplay between mutation, gene conversion, unequal crossover events, and sexual reproduction increases genetic divergence between the ITS regions of different species while maintaining a lower than expected within-species variation (Gerbi 1986, Silva and Faust 1995, Giacobbe and Yang 1999, Probert et al. 2002, Lee et al. 2003, Nagai et al. 2003, Shankle et al. 2004). This implies that if gene conversion is operating effectively, intraspecific ITS region genetic distances would be consistently lower than interspecific distances. A threshold genetic distance may therefore exist that correlates with known species-level differences.

MATERIALS AND METHODS

Calculating genetic distances. Sequences were aligned using the CLUSTAL_X algorithm (Thompson et al. 1997). The open and extended gap penalties were set at the default values of 10.0 and 5.0, respectively, for both the pair-wise and multiple alignment phases. Completed alignments were saved as NEXUS files and imported into PAUP*4b10 software (Swofford 2002) so that divergence rates could be estimated using simple uncorrected pair-wise (p) distance matrices. Previous studies have indicated that the ITS1 and ITS2 regions exhibit different levels of variation, which could influence the overall genetic distances estimated within or between species. To investigate this possibility, p values were calculated for the ITS1, 5.8S, and ITS2 regions separately, as well as for the combined ITS region sequence data. Genetic distances between sequences in each analysis were expressed as the number of substitutions per site. Constant alignment parameters and a simple distance method make the procedure easy to replicate.

Sampling to assess within-genome ITS region genetic distances. The data for assessing within-genome variability among ITS region copies came primarily from two sources. The first source included ITS copies obtained from single-cell isolates that were sequenced as a prerequisite for developing the species-specific ITS PCR assays reported in Litaker et al. (2003). It was assumed that because the cultures were started from a single cell, all the ITS region copies sequenced from that isolate belong to the same genome. The specific dino-

flagellates sequenced included the following: (1) *Amylodinium ocellatum* E. Brown et Hovasse; (2) *Cryptoperidiniopsis brodyi* Steid., Landsberg, P. L. Mason, Vogelbein, Tester et Litaker isolate Cell H; (3) *Cryptoperidiniopsis* sp. isolate A5 (= PLO21) (Steidinger et al. 2006); (4) *C. brodyi* isolate V14; (5) *Heterocapsa triquetra* (Ehreb.) F. Stein; (6) *Karenia brevis* (C. C. Davis) G. Hansen et Moestrup; (7) *K. mikimotoi* (Miyake et Kominami ex Oda) G. Hansen et Moestrup; (8) *Karlodinium veneficum* (D. L. Ballant.) J. Larsen; (9) a Lucy species (Steidinger et al. 2001); (10) *Pfiesteria piscicida* Steid. et Burkh. J. M. isolate CCMP 1834; (11) *Pf. piscicida* isolate CCMP 1921; (12) *Pf. piscicida* isolate Noga P; (13) *P. piscicida* isolate Cell M; (14) *Pf. piscicida* isolate VIMS 10/11; (15) *Pseudopfiesteria shumwayae* (H. B. Glasgow et J. M. Burkh.) Litaker, Steid., P. L. Mason, J. H. Shields et Tester isolate VIMS 1049; (16) *Ps. shumwayae* isolate Noga S; and (17) *Prorocentrum minimum* (Pavill.) J. Schiller. To ensure consistent results, single-cell isolates were also established from all CCMP cultures. The second source included a large study of intragenomic variation in the ITS region of *Scrippsiella* species (Montresor et al. 2003). The corresponding GenBank accession numbers for these species are listed in the supplementary materials. Depending on the study, the ITS regions were PCR-amplified from either individual cells or single-cell isolates and then subcloned and sequenced. Sequence differences among the PCR clones from each isolate were then used to estimate within-genome genetic distances. The PCR-amplification, cloning, and sequencing procedures used in these studies, as well as information on the culture conditions, are given in Litaker et al. (2003) and Montresor et al. (2003).

Sampling to assess intergenomic genetic distances. Here the term *intergenomic* will only be used to refer to ITS region copies obtained from different individuals of the same species and not to those obtained from different species. Intergenomic distances were calculated separately for ITS1, 5.8S and ITS2, as well as for the combined ITS region. For these analyses, ITS region copies belonging to individuals of 35 different species were assembled from GenBank. Sequences were included only if they originated from independent single-cell isolates of the same putative species. It should be noted that the species designations assigned to some sequences are likely incorrect, with the same sequences being attributed to more than one species. In these cases, the sequences were combined in the same analysis, and the ambiguous sequence assignments noted. The references used to define the genus and species identities in this study, including cryptic or recently recognized species, were as follows: (1) *Alexandrium tamarense* (Lebour) Balech and variants (Adachi et al. 1996, Sako 1999); (2) *Dinophysis norvegica* Ehrenb. and the combination of *D. accuminata* Claparède et Lachmann/*dens* Pavillard/*parvillardii* Schröder and *sacculus* F. Stein as conspecific, with *D. acuta* Ehrenb. being difficult to delineate (Zingone et al. 1998, Giacobbe et al. 2000, Marin et al. 2001, Edvardsen et al. 2003); (3) split of *Karenia*/*Karlodinium* (Daugbjerg et al. 2000); (4) *Ostreopsis ovata* Fukuyo variants (Pin et al. 2001); and (5) the cryptic *Scrippsiella* species (D'Onofrio et al. 1999, Montresor et al. 2003). The specific GenBank numbers used in the analysis are listed in the supplementary materials. The combination of intergenomic and within-genome variation will be referred to as "within-species" variation.

Sampling to assess between-species ITS region genetic distances. The ITS region sequence data from 81 species belonging to 14 separate genera were obtained from GenBank. Sequence divergence between species belonging to different genera was extensive for the ITS region and resulted in too much ambiguity in the alignments for reliable comparisons to be made. Intraspecific genetic distances were therefore estimated from ITS region alignments of the species within each individual genus. In cases where multiple ITS region sequences were available for a given species, the most frequently observed sequence was chosen to represent that species. This selection process was adopted after all the sequences for several species in two genera were compared in sequential analyses and determined to produce nearly identical genetic distance estimates. Intergenomic distances were calculated separately for ITS1, 5.8S, and ITS2, as well as for the combined ITS1/5.8S/ITS2 region, as described previously. The GenBank sequences used to estimate the between-species genetic distances are listed in the supplementary material.

Estimating the boundaries between the 5.8S gene and flanking ITS regions. To calculate differences in divergence estimates between ITS1, the 5.8S gene, and ITS2, it was necessary to delineate a consistent boundary between the 5.8S gene and the flanking ITS regions. In this study, the 5.8S gene boundaries were defined using Nazar's (1984) 5.8S alignment data for an evolutionarily diverse array of organisms and the nuclease S1 mapping of the *Prorocentrum minimum* 5.8S RNA (Maroteaux et al. 1985). Using these data, the 5′ boundary of the 5.8S gene is defined as 5′-ACA(A/g/t)(C/T/g)(a/g/T)T(g/T)C(A/g)G(C/T)(a/G)(A/g)(C/T)(c/G)(a/G)AT-3′, and the corresponding 3′ boundary as 5′-G(A/g)(A/c/G)(A/g)(a/c/G)(C/T)(A/g)(c/g/T)(a/g/T) (C/T)(a/C/T)(c/T)(a/G/t)C(c/g/T)T(C/t)(A/c/g)G(a/T)G(C/T)(C/T)(a/c/T)(A/c/g/t)(A/C/G/T)-3′. Nucleotides within parentheses represent each of the bases observed at that particular position in the alignment. Bases not in parentheses represent nucleotide positions conserved among all species. Nucleotides shown in lowercase indicate that the particular base was rarely observed at that position. For example, the designation (A/c/g) indicates that most of the aligned sequences contained an A at that nucleotide position, with only a relatively few having either C or G.

RESULTS

Within-genome variation. Sequencing of multiple ITS region copies from individual genomes revealed that the number of observed polymorphic sites ranged between 2 and 27 per genome depending on the species (Fig. 1). These polymorphisms were distributed among different alleles, and each variant generally differed from all other alleles by only a few bp. Most of the observed polymorphisms were simple nucleotide substitutions with few insertion/deletion events. Approximately 75% of these intragenomic mutations were transitions. A majority of the polymorphic sites occurred in the ITS1 and ITS2 regions, with fewer polymorphic sites observed in the 5.8S gene sequence (Figs. 1 and 2).

To obtain a preliminary understanding of the abundance of variant ITS region alleles and their distribution among genomes of the same species, multiple ITS region copies were sequenced from two isolates of *Pseudopfiesteria shumwayae* and four isolates of *Pf. piscicida*. In each species, there was a common ITS region sequence recovered from a majority of the DNA clones. Each genome also exhibited a number of less-frequent variant copies (Fig. 2). Some of these variants were observed from more than one isolate of the same species.

Intergenomic genetic distances. The intergenomic ITS region genetic distances measured between individuals belonging to the same species varied between 0 and 0.0206 substitutions per site and were similar to the intragenomic variation, which ranged from 0 to 0.0170

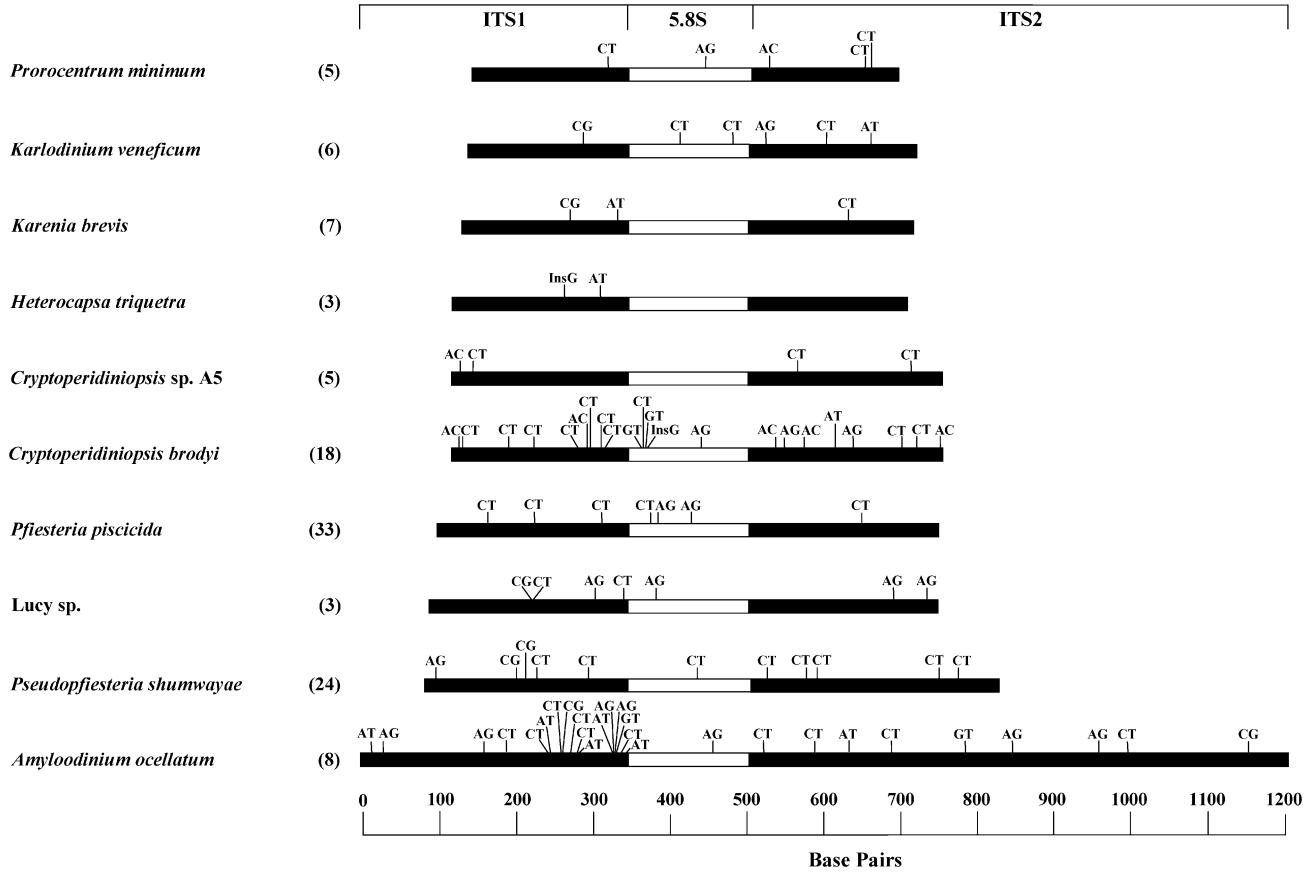


FIG. 1. Within-genome variation of ITS region copies from 10 representative dinoflagellate species. The numbers in parentheses indicate the number of copies sequenced. Nucleotide pairs indicate multiple nucleotides (nucleotide polymorphisms) were found at that site in the various copies sequenced. The abbreviation "Ins" followed by a nucleotide indicates a single base pair insertion at that point in the sequence. The species are arranged in order of increasing size of the combined ITS1/5.8S/ITS2 regions to illustrate the variation in the length of the ITS1 and ITS2 regions observed in representative dinoflagellate species. *Cryptoperidiniopsis brodyi* was recently described (Steidinger et al. 2006), and the data presented for this species were derived from isolates H and V14.

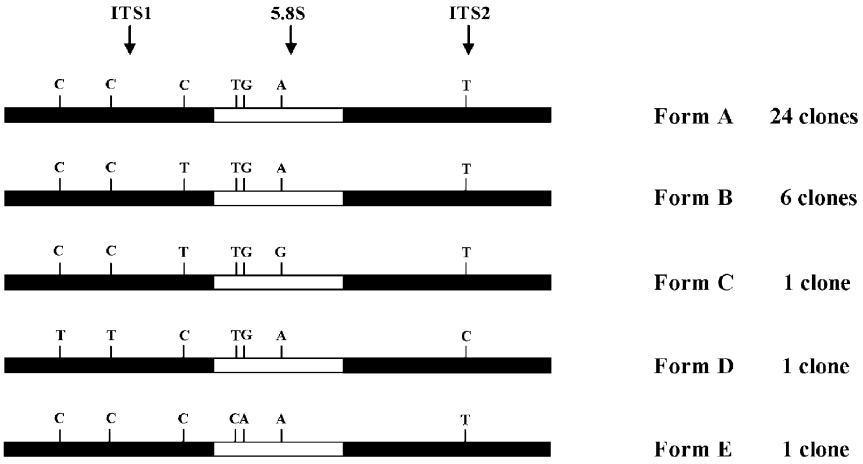
substitutions per site (Table 1). The corresponding intergenomic p values calculated for just the ITS1 region ranged from 0 to 0.0153, and from 0 to 0.0207 for ITS2. These ITS distances were significantly higher than could be accounted for by the known incorporation error rates of either the *Taq* ($\sim 5 \times 10^{-4}$ errors per bp) or Platinum *Taq* High Fidelity ($\sim 5 \times 10^{-5}$ errors per bp) DNA polymerases used in these studies (Tindall and Kunkel 1988, Cariello et al. 1991).

Between-species genetic distances. The ITS region p values between species in the same genera ranged from 0.042 to 0.580. The lowest of these between-species genetic distances was higher than any p values observed within a species ($p \leq 0.0206$; Fig. 3A). Between-species p values calculated using only aligned ITS1 regions ranged from 0.0303 to 0.734 and slightly overlapped the highest observed within-species range of 0–0.0344 (Fig. 3B, Table 1). This overlap was caused by the high variation among ITS region sequences within *A. ocellatum* and the relatively high similarity observed in ITS1 for the species pair *Symbiodinium pilosum* Trench et Blank and *S. muscatinei* LaJeunesse et Trench. The remaining between-species ITS1 genetic distances were consistently higher ($p = 0.0550$–0.734) than those observed within a species (Fig. 3B; Table 1).

The between-species genetic distances calculated for alignments of the ITS2 region alone ($p = 0.0210$–0.732) were comparable to those observed for the ITS1 region. The between-species variation was generally much higher than the corresponding within-species variation, with the exception of the similarity between the ITS2 sequence of a single species pair, *Scrippsiella rotunda* J. Lewis and *S. operosa* (Deflandre) Montresor (Fig. 3D). All other interspecific ITS2 genetic distances exceeded $p = 0.0354$ substitutions per site, with most having p values > 0.0706 (Table 1).

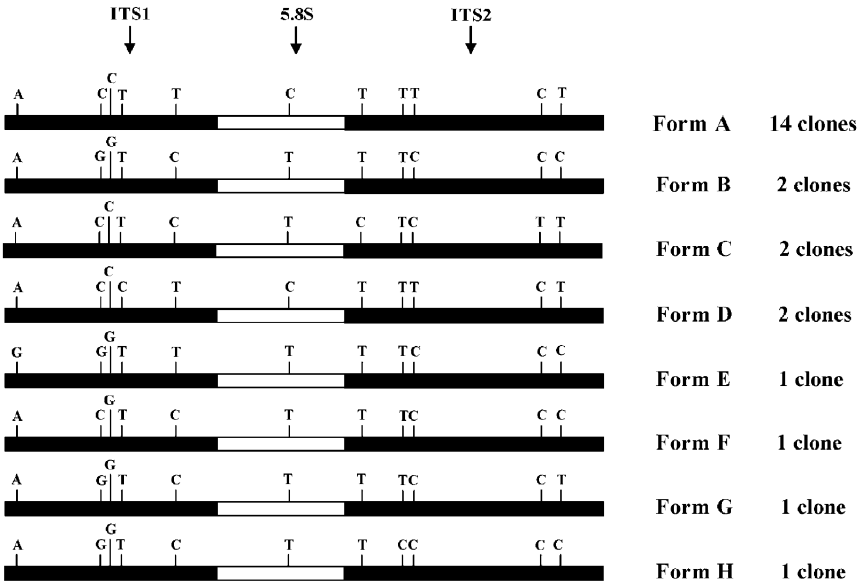
The between-species genetic distances in the 5.8S region were generally lower between species than distances observed in the corresponding ITS1 and ITS2 regions (Fig. 3C). These lower genetic distances resulted in a much greater degree of overlap in the distances within and between species. For a majority of species, the mean p value between species within genera was substantially higher than that observed within a single species (Fig. 3C).

Within-genome nucleotide variation in the ITS region of *Pfiesteria piscicida*



Cell isolate (Total clones sequenced)	CCMP 1921 (12)	Noga P (12)	CCMP 1834 (7)	VIMS10/11 (2)
ITS forms found in each isolate. The labels A – E correspond to the diagram shown above.	A B C	A D E	A B	A

Within-genome nucleotide variation in the ITS region of *Pseudopfiesteria shumwayae*



Cell isolate (Total clones sequenced)	VIMS 1049 (12)	Noga S (12)
ITS forms found in each isolate. The labels A – H correspond with the diagram shown above.	A B C D E	A F G H

FIG. 2. Intragenomic nucleotide variation in the ITS region of two dinoflagellate species based on multiple single-cell isolates. This analysis was undertaken to obtain a preliminary estimate of the abundance of variant ITS1/5.8S/ITS2 copies distributed among genomes of the same species. Five different nucleotide sequences (Forms A–E) were obtained in the four *Pfiesteria piscicida* genomes examined. Form A was the most frequent across all four genomes, representing 24 of the 33 clones sequenced. The next most abundant form, Form B, was found in two of the four genomes. The remaining three forms were unique to only one genome. Eight different copies (Forms A–H) were observed in two *Pseudopfiesteria shumwayae* genomes. Form A was dominant, accounting for 14 of the 24 total sequences. Variant alleles were more frequent in *Ps. shumwayae* than in *Pf. piscicida*.

DISCUSSION

Recognizing putative dinoflagellate species using ITS region genetic distances. The primary objective of this study was to evaluate whether uncorrected genetic distances (p) between ITS region rDNA sequences would correlate with species boundaries that were based on morphological criteria. The hypothesis was that if ITS region divergences between species were consistently high relative to the variation observed within species, then genetic distances could be used to recognize putative species. These putative species could in turn serve as a basis for formal taxonomic evaluation or for testable hypotheses concerning the diversity, distribution, and toxicity of certain dinoflagellates. Results from 81 species in 14 genera showed that between-species p values varied from 0.042 to 0.580 substitutions per site, and that most distances between species exceeded 0.08. In contrast, the intraspecific genetic distances calculated for 35 species were uniformly lower, ranging from 0 to 0.0206 (Figs. 1–2; Table 1). Based on these data, ITS region distances ≥ 0.04 are indicative of species-level divergences.

While not observed in this study, the possibility exists that rates of evolution may be significantly higher or lower in certain genera. This implies that in some cases, the p value used to recognize species may have to be recalculated. The current ITS region data suggested that appropriate p values for recognizing species in these genera could be established by determining the average within- and between-species variation. In instances where accelerated rates of evolution have occurred, the $p \geq 0.04$ threshold would likely still hold, provided the within-species divergences remained in the range observed in this study (Table 1).

Circumstances where divergences are low due to slower rates of evolution, or due to recent speciation, would prove more difficult to resolve. In these instances, p values between species would likely be <0.04. For example, *D. acuta* and *D. acuminata* exhibit similar morphologies and can vary greatly in size, making it difficult to identify them (Giacobbe et al. 2000, Marin et al. 2001, Edvardsen et al. 2003). The existing morphological and molecular data for these two species are contradictory. The ITS region p value <0.04 indicates they are the same species. In contrast, slight morphological differences and the available SSU data indicate that they are different species. The LSU data are inconclusive (Edvardsen et al. 2003). Taken together, these data indicate that *D. acuta* and *D. acuminata* are two recently diverged species that cannot be identified using ITS data alone.

The genus where divergence rates are likely to be most problematic for calculating p-value thresholds is *Symbiodinium*. *Symbiodinium* species are morphologically indistinct and demonstrate a remarkable ability to form symbiotic relationships with numerous host species, including cnidarians, mollusks, flatworms, and protozoa. Formation of new symbiotic relationships in many cases confers genetic isolation, as well as a suite of unique selection pressures depending on the environmental conditions within the host's habitat. There is evidence that this isolation gives rise to new species with ITS genetic divergences <0.04 (van Oppen et al. 2001, LaJeunesse et al. 2003, 2004). Similar diversification is less likely for free-living species, which are not generally subject to selective forces that would confer the same degree of genetic isolation.

An alternative method for resolving species identification in cases where there are small genetic distances is to sequence other genes as independent tests. Good candidate genes include the SSU and LSU genes that flank the ITS regions. These genes can be PCR-amplified and sequenced in conjunction with the ITS region and contain variable domains that are phylogenetically informative (Scholin et al. 1995, Litaker et al. 1999, Jorgensen et al. 2004, Ordas et al. 2004, Saldarriaga et al. 2004). Alternatively, phylogenetic analyses based on single copy genes, chloroplast, or mitochondrial genes that typically exhibit less intragenomic variation could be used to resolve ITS ambiguities (Zhang et al. 2000, Zhang and Lin 2002). In order to support species-level differences, these phylogenetic analyses would have to demonstrate the putative species as being monophyletic.

Interpreting the results of shotgun cloning studies. The results of this study have implications for how small bp changes among rDNA sequences derived from shotgun cloning studies are interpreted. In shotgun cloning studies, DNA from environmental samples is extracted, PCR-amplified with conserved primers, cloned, and sequenced to assess community diversity (Venter et al. 2004). Most commonly, rDNA regions are the target sequence (Rappe et al. 2000, Suzuki et al. 2001, Moreira and Lopez-Garcia 2002). The result is numerous sequences with varying degrees of similarity, depending upon the diversity of organisms present in the samples. In the case of ITS sequences, as well as other rDNA regions, dominant and variant rDNA copies will be randomly sequenced. If the genetic distance used for distinguishing potential species is set too low, many of the variant rDNA copies may be incorrectly interpreted as distinct species. Conversely, if the criterion is set too high, the actual number of species will be underestimated.

Potential complications posed by pseudogenes. The existence of rDNA pseudogenes must also be considered when evaluating rDNA sequence variation. A pseudogene is a remnant copy of a once-functional gene that has accumulated sufficient mutations to render it nonfunctional. The existence of pseudogenes is well known for dinoflagellates (Scholin et al. 1993, Santos et al. 2003). If these pseudogenes have diverged sufficiently, they could be incorrectly interpreted as sequences from a different species. Pseudogenes can often be identified by mutations in conserved coding regions (Scholin et al. 1993, Scholin and Anderson 1994) or by significant deletions, insertions, or structural rearrangements in the gene sequence.

TABLE 1. Estimated genetic distances (p) between ITS regions from the same genome, between individuals of the same species, and between species. Substitutions per site were calculated for the combined ITS region as well as for the separate ITS1, 5.8S, and ITS2 regions.

			Combined ITS1, 5.8S, ITS2			ITS1			5.8S			ITS2		
Taxa		n	Mean	Maximum	Minimum	Mean	Maximum	Minimum	Mean	Maximum	Minimum	Mean	Maximum	Minimum
Within genome														
1	*Amyloodinium ocellatum*	$n = 8$	0.0114	0.0170	0.0062	0.0241	0.0334	0.0086	0.0031	0.0063	0.0000	0.0060	0.0090	0.0038
2	*Cryptoperidiniopsis brodyi* CellH	$n = 3$	0.0026	0.0040	0.0013	0.0026	0.0038	0.0000	0.0000	0.0000	0.0000	0.0039	0.0060	0.0026
3	*Cryptoperidiniopsis* species A5	$n = 5$	0.0035	0.0055	0.0014	0.0043	0.0087	0.0000	0.0000	0.0000	0.0000	0.0045	0.0060	0.0030
4	*Cryptoperidiniopsis brodyi* V14	$n = 18$	0.0090	0.0135	0.0014	0.0128	0.0192	0.0039	0.0083	0.0125	0.0000	0.0063	0.0094	0.0000
5	*Heterocapsa triquetra / pygmaea*	$n = 3$[a,b]	0.0017			0.0043						0.0000		
6	*Karenia brevis*	$n = 7$[a]	0.0045			0.0063						0.0034		
7	*Karenia mikimotoi*	$n = 2$	0.0000			0.0000			0.0000			0.0000		
8	*Karlodinium veneficum*	$n = 6$	0.0070	0.0089	0.0015	0.0032	0.0047	0.0000	0.0084	0.0126	0.0063	0.0089	0.0133	0.0000
9	Lucy sp.	$n = 3$	0.0063	0.0080	0.0027	0.0103	0.0115	0.0039	0.0042	0.0063	0.0000	0.0041	0.0061	0.0031
10	*Pfiesteria piscicida* CCMP 1834	$n = 7$	0.0046	0.0082	0.0014	0.0056	0.0120	0.0000	0.0075	0.0188	0.0000	0.0025	0.0062	0.0000
11	*Pfiesteria piscicida* CCMP 1921	$n = 12$	0.0020	0.0031	0.0014	0.0027	0.0040	0.0000	0.0042	0.0063	0.0000	0.0000	0.0000	0.0000
12	*Pfiesteria piscicida* Noga P	$n = 12$	0.0054	0.0082	0.0041	0.0053	0.0080	0.0000	0.0010	0.0125	0.0000	0.0041	0.0062	0.0031
13	*Pfiesteria piscicida* Cell M	$n = 2$	0.0014			0.0040			0.0000			0.0000		
14	*Pfiesteria piscicida* VIMS10/11	$n = 2$	0.0000			0.0000			0.0000			0.0000		
15	*Pseudopfiesteria shumwayae* 1049	$n = 12$	0.0046	0.0084	0.0000	0.0042	0.0112	0.0000	0.0026	0.0063	0.0000	0.0034	0.0124	0.0000
16	*Pseudopfiesteria shumwayae* Noga S	$n = 12$	0.0035	0.0096	0.0000	0.0043	0.0096	0.0000	0.0044	0.0096	0.0000	0.0084	0.0096	0.006
17	*Prorocentrum minimum*	$n = 5$[a]	0.0054			0.0049			0.0063			0.0053		
18	*Scrippsiella trochoidea* SN33	$n = 2$	0.0037			0.0000			0.0063			0.0052		
19	*Scrippsiella trochoidea* SN61	$n = 3$	0.0085	0.0110	0.0036	0.0034	0.0052	0.0000	0.0000	0.0000	0.0000	0.0205	0.0256	0.0103
20	*Scrippsiella trochoidea* SZN64	$n = 3$	0.0036	0.0055	0.0018	0.0034	0.0051	0.0000	0.0000	0.0000	0.0000	0.0069	0.0103	0.0000
21	*Scrippsiella trochoidea* SZN69	$n = 3$	0.0061	0.0091	0.0018	0.0068	0.0102	0.0000	0.0000	0.0000	0.0000	0.0103	0.0156	0.0052
22	*Scrippsiella trochoidea* SZN70	$n = 5$	0.0055	0.0091	0.0018	0.0061	0.0102	0.0000	0.0000	0.0000	0.0000	0.0093	0.0206	0.0000
23	*Scrippsiella trochoidea* SZN72	$n = 3$	0.0036	0.0055	0.0000	0.0000	0.0000	0.0000	0.0042	0.0063	0.0000	0.0069	0.0103	0.0000
24	*Scrippsiella trochoidea* SZN76	$n = 4$	0.0049	0.0073	0.0018	0.0105	0.0158	0.0000	0.0000	0.0000	0.0000	0.0034	0.0051	0.0000
25	*Scrippsiella trochoidea* SZN77	$n = 4$	0.0081	0.0148	0.0018	0.0173	0.0339	0.0048	0.0025	0.0063	0.0000	0.0025	0.0063	0.0000
26	*Scrippsiella trochoidea* SZN89	$n = 7$	0.0058	0.0110	0.0000	0.0029	0.0103	0.0000	0.0108	0.0252	0.0000	0.0044	0.0104	0.0000
27	*Scrippsiella trochoidea* SZN91	$n = 4$	0.0046	0.0091	0.0000	0.0051	0.0102	0.0000	0.0000	0.0000	0.0000	0.0078	0.0156	0.0000
28	*Scrippsiella trochoidea* SZN93	$n = 2$	0.0000			0.0000			0.0000			0.0000		
29	*Scrippsiella trochoidea* var *aciculifera* SZN60	$n = 3$	0.0036	0.0055	0.0018	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0102	0.0153	0.0051
30	*Scrippsiella trochoidea* var *aciculifera* SZN63	$n = 2$	0.0000			0.0000			0.0000			0.0000		
Between individuals of same species														
1	*Alexandrium affine*	$n = 3$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	*Alexandrium catenella*	$n = 4$	0.0049	0.0097	0.0000	0.0060	0.0121	0.0000	0.0063	0.0126	0.0000	0.0026	0.0053	0.0000
3	*Alexandrium lusitanicum / minutum*	$n = 4$	0.0000			0.0000			0.0000			0.0000		
4	*Alexandrium pseudogonyaulax* AS-1	$n = 2$	0.0059			0.0000			0.0187			0.0000		
5	*Alexandrium tamarense / fundyense* gene A	$n = 6$	0.0013	0.0039	0.0000	0.0020	0.0060	0.0000	0.0000	0.0000	0.0000	0.0017	0.0051	0.0000
6	*Alexandrium tamarense / fundyense* gene B	$n = 2$	0.0019			0.0060			0.0000			0.0000		
7	*Alexandrium tamarense* WKS-1 type	$n = 2$	0.0019			0.0000			0.0063			0.0000		
8	*Alexandrium taylori*	$n = 5$	0.0008	0.0020	0.0000	0.0000	0.0000	0.0000	0.0025	0.0063	0.0000	0.0000	0.0000	0.0000
9	*Amphidinium carterae*	$n = 2$	0.0046			0.0105			0.0000			0.0000		
10	*Amphidinium eilantiensis*	$n = 2$	0.0000			0.0000			0.0000			0.0000		
11	*Coolia malayensis*	$n = 8$	0.0049	0.0133	0.0000	0.0105	0.0400	0.0000	0.0050	0.0126	0.0000	0.0000	0.0000	0.0000
12	*Coolia monotis*	$n = 2$	0.0027			0.0000			0.0000			0.0096		
13	*Dinophysis acuta, acuminata, dens, pavillardii, sacculus*	$n = 19$	0.0033	0.0121	0.0000	0.0032	0.0099	0.0000	0.0032	0.0099	0.0000	0.0050	0.0186	0.0000
14	*Dinophysis norvegica*	$n = 3$	0.0006	0.0018	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0018	0.0054	0.0000
15	*Gymnodinium catenatum*	$n = 20$	0.0011	0.0077	0.0000	0.0004	0.0059	0.0000	0.0013	0.0121	0.0000	0.0015	0.0109	0.0000
16	*Gymnodinium corii / simplex*	$n = 2$	0.0000			0.0000			0.0000			0.0000		

17	*Gymnodinium linucheae*	$n = 4$	0.0018	0.0036	0.0000	0.0025	0.0049	0.0000	0.0000	0.0000	0.0000	0.0027	0.0054	0.0000
18	*Gymnodinium* sp. zhao01/zhao08	$n = 3$	0.0012	0.0018	0.0000	0.0000	0.0000	0.0000	0.0042	0.0063	0.0000	0.0000	0.0000	0.0000
19	*Gyrodinium instriatum*	$n = 4$	0.0042	0.0066	0.0000	0.0021	0.0043	0.0000	0.0000	0.0000	0.0000	0.0096	0.0191	0.0000
20	*Heterocapsa circularisquama*	$n = 4$	0.0000			0.0000			0.0000			0.0000		
21	*Heterocapsa pygmaea / triquetra*	$n = 2$	0.0000			0.0000			0.0000			0.0000		
22	*Heterocapsa triquetra*	$n = 3$	0.0000			0.0000			0.0000			0.0000		
23	*Heterocapsa* FK6 NIES614	$n = 2$	0.0017			0.0030			0.0000			0.0000		
24	*Karenia mikimotoi*	$n = 2$	0.0000			0.0000			0.0000			0.0000		
25	*Ostreopsis ovata* Malacca Straights	$n = 2$	0.0000			0.0000			0.0000					
26	*Ostreopsis ovata* South China Sea	$n = 9$	0.0139	0.0206	0.0000	0.0082	0.0211	0.0000	0.0036	0.0065	0.0000			
27	*Ostreopsis* sp. CSIC D5 cf. *siamensis*	$n = 2$	0.0000			0.0000			0.0000			0.0000		
28	*Pfiesteria piscicida*	$n = 26$	0.0006	0.0068	0.0000	0.0012	0.0120	0.0000	0.0005	0.0063	0.0000	0.0002	0.0031	0.0000
29	*Pseudopfiesteria shumwayae*	$n = 2$	0.0084			0.0038			0.0000			0.0000		
30	*Prorocentrum micans*	$n = 3$	0.0072	0.0090	0.0054	0.0065	0.0097	0.0049	0.0042	0.0063	0.0000	0.0106	0.0159	0.0053
31	*Prorocentrum minumum*	$n = 2$	0.0072			0.0097			0.0000			0.0106		
32	*Scrippsiella trochoidea* SZN33,89,90	$n = 16$	0.0056	0.0147	0.0000	0.0039	0.0103	0.0000	0.0069	0.0253	0.0000	0.0063	0.0207	0.0000
33	*Scrippsiella trochoidea* SZN63,60	$n = 5$	0.0022	0.0055	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0061	0.0153	0.0000
34	*Scrippsiella trochoidea* SZN64,69,70,72,91,93	$n = 20$	0.0040	0.0128	0.0000	0.0038	0.0153	0.0000	0.0007	0.0063	0.0000	0.0069	0.0207	0.0000
35	*Symbiodinium microadriaticum*	$n = 2$	0.0037			0.0000			0.0065			0.0056		
Between species in the same genus														
1	*Alexandrium*	$n = 13$	0.3771	0.5211	0.0424	0.5100	0.6719	0.1097	0.1401	0.2812	0.0127	0.4738	0.7184	0.0354
2	*Amphidinium*	$n = 2$	0.1641			0.2624			0.0192			0.2251		
3	*Coolia*	$n = 2$	0.2340			0.3334			0.0703			0.3831		
4	*Cryptoperidiniopsis*	$n = 2$	0.1801			0.1147			0.0917			0.2638		
5	*Dinophysis*	$n = 5$	0.2033	0.3497	0.0443	0.2453	0.4283	0.0550	0.0781	0.1438	0.025	0.2964	0.5113	0.0448
6	*Gonyaulax*	$n = 2$	0.6263			0.5857			0.5671			0.6947		
7	*Gymnodinium*	$n = 8$	0.4868	0.5768	0.2965	0.6351	0.7341	0.5195	0.1544	0.2516	0.0323	0.5954	0.7321	0.2815
8	*Heterocapsa*	$n = 9$	0.1896	0.2685	0.0855	0.2347	0.3897	0.0927	0.0555	0.0950	0.0150	0.3166	0.3887	0.1532
9	*Karenia*	$n = 2$	0.0780			0.0573			0.0627			0.0975		
10	*Ostreopsis*	$n = 5$	0.2496	0.3586	0.0816	0.3634	0.4785	0.1362	0.1487	0.2738	0.0130	0.3179	0.4546	0.1230
11	*Pentapharsodinium*	$n = 2$	0.0817			0.1020			0.0251			0.1073		
12	*Prorocentrum*	$n = 4$	0.2126	0.2551	0.1072	0.1657	0.2045	0.0717	0.0328	0.0492	0.0164	0.2983	0.3689	0.1272
13	*Scrippsiella*	$n = 13$	0.1533	0.2343	0.0584	0.1766	0.2709	0.0728	0.0334	0.1163	0.0000	0.2382	0.3500	0.0198
14	*Symbiodinium*	$n = 7$	0.3505	0.4794	0.0487	0.3187	0.4469	0.0379	0.0891	0.1344	0.0000	0.4736	0.6500	0.0735

For the "Within genome" data, n = the number of PCR clones sequenced from each genome. For the "Between individuals of same species" data, n = the number of single-cell isolates from which ITS loci were sequenced. A list of sequences used to calculate the mean, maximum, and minimum divergence values is presented in the supplementary materials. Note: Many of the PCR clones sequenced were identical, either within a genome or between individuals of the same species. These identical sequences were not entered into GenBank. As a result, n values in this chart, which include the redundant sequences, are often greater than the number of unique GenBank accession numbers listed in the supplementary materials.

[a]More than two clones were sequenced, but only two unique sequences were obtained.

[b]Identical sequences have been attributed to two different species. A mean of 0 indicates that all the sequences were identical.

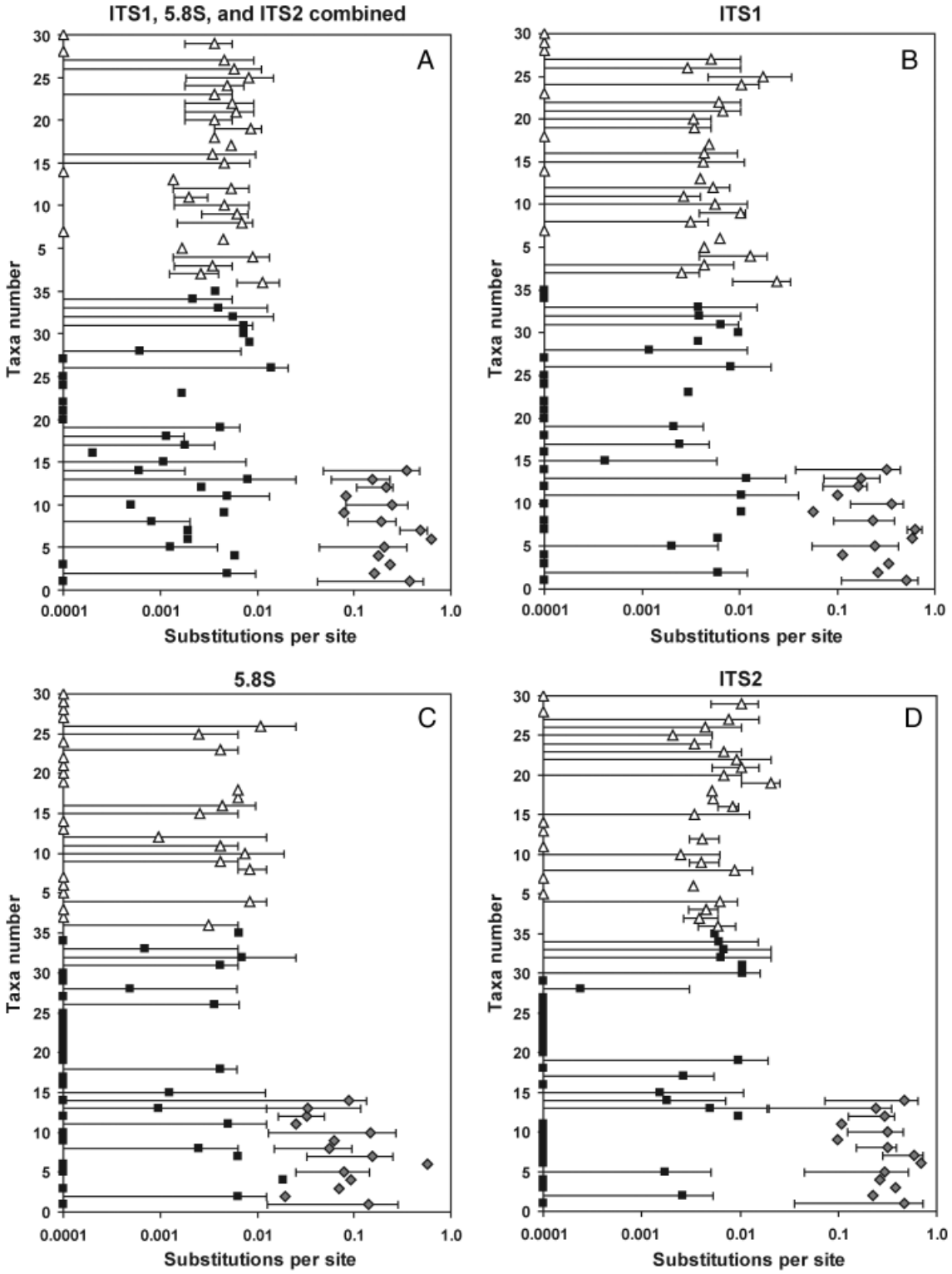


FIG. 3. Plot of the mean, maximum, and minimum substitutions per site calculated for ITS copies from the same genome (Δ), among individual ITS copies from different individuals belonging to the same species (■), and among species belonging to the same genus (◇). The sequences for each group analyzed were aligned, and substitutions per site were estimated for the individual ITS1, 5.8S, and ITS2 regions as well as for combined ITS region sequences. The error bars indicate the maximum and minimum substitutions per site measured. In instances where all the sequences were identical, the substitutions per site were listed as 0. Each of the 0 values was arbitrarily plotted on the log scale for the *x*-axis as 0.0001 substitutions per site. The sequences used in calculating the various genetic distances are listed in the supplementary material.

TABLE 2. Examples from five genera demonstrating inconsistencies between-species identification based on morphology and on ITS sequences.

Species	*GenBank accession number*	*Species*	*GenBank accession number*
Alexandrium		*Dinophysis* grouping continued	
Alexandrium lusitanicum 181NT	AY455825	*Dinophysis sacculus* from France, Urbino Lagoon	AY040581
Alexandrium lusitanicum	AY455826	*Dinophysis sacculus* from France, Bay of Toulon	AY040582
Alexandrium minutum	AJ318460	*Dinophysis sacculus* from France, Urbino Lagoon	AY040583
		Dinophysis sacculus	AJ012007
Alexandrium tamarense strain WKS-1	AB006991	*Dinophysis sacculus* from Spain	AJ304807
Alexandrium tamarense	AJ005047		
Alexandrium tamarense isolate CCMP1493	AJ005048	*Gymnodinium*	
Alexandrium catenella	AY347308	*Gymnodinium linucheae*	AF333509
Alexandrium catenella isolate 4	AJ298900	*Symbiodinium* sp. clade A	AF427465
Alexandrium catanella strain MI7	AB006990	*Symbiodinium* isolate 47-5i	AY074984
		Symbiodinium isolate 47-5iii	AY074985
Alexandrium fundyense	AJ005049		
Alexandrium tamarense B strain FK-788	AB006994	*Gymnodinium* sp. zhao08	AJ534384
Alexandrium tamarense A strain FK-788	AB006993	*Gymnodinium* sp. zhao 01	AJ534385
Alexandrium tamarense isolate AT-A	AF374224	*Symbiodinium* sp. G15	AY160123
Alexandrium tamarense isolate AT-B	AF374225		
Alexandrium tamarense isolate AT-10	AF374226		
Alexandrium tamarense isolate AT-2	AF374227	*Heterocapsa*	
Alexandrium tamarense isolate AT-6	AF374228	*Heterocapsa pygmaea* CCMP1490	AB084094
		Heterocapsa triquetra CCMP 448	AF352363
Dinophysis		*Heterocapsa triquetra* CCMP 448	AF352364
Dinophysis acuminata from Spain	AJ272119	*Heterocapsa pygmaea* CCMP 1322	AB084093
Dinophysis acuminata isolate DacmF4-9	AJ506986		
Dinophysis acuminata from Australia	AY040573	*Heterocapsa triquetra* isolate NIES 7	AB084101
Dinophysis acuminata from United Kingdom	AY040574	*Heterocapsa triquetra* strain HT-1	AF208249
Dinophysis acuminata from France	AY040575	*Heterocapsa triquetra* CCMP448	AF527816
Dinophysis acuminata from France	AY040576		
Dinophysis acuminata from France	AY040579	*Heterocapsa* sp. FK6-D47	AB084097
Dinophysis acuminata from France	AY040577	*Heterocapsa* sp. NIES614	AB084099
Dinophysis acuminata from France	AY040578		
Dinophysis dens from Portugal	AY040571		
Dinophysis dens from France	AY040572	*Ostreopsis*	
Dinophysis pavillardii	AJ404000	*Ostreopsis* cf. *siamensis* CNR-B4	AJ301643
Dinophysis sacculus from France, Urbino Lagoon	AY040580	*Ostreopsis* sp. CSIC-D5	AJ312944

Groupings in the table indicate that all members of the group have identical or nearly identical sequences that are consistent with the observed within-species variation, even though one or more members of the group had been identified morphologically as a different species or strain.

DNA barcoding of dinoflagellate species. In 2004 a consortium of major natural history museums and herbaria launched the ''Barcode of Life Initiative.'' The goal of this project was to encourage an ''emerging collaborative effort to promote a process enabling the rapid low cost identification of the estimated 10 million species of Earth's fauna and flora'' (Biological Conservation Newsletter 2004). The feasibility of using a barcoding system has been validated using an approximately 700 bp region of the mitochondrial cytochrome *c* oxidase subunit I (COI) gene (Hebert et al. 2003, 2004, Hogg and Hebert 2004). Takabayashi et al. (2004) recently completed a COI-based phylogeny for the major *Symbiodinium* groups that was congruent with those inferred from chloroplast LSU and nuclear ribosomal genes. Data on additional COI sequences, however, suggest that COI genes may prove problematic for identification of some dinoflagellate species due to complex structure and that cytochrome *b* (COB) may be a better choice (S. J. Lin, personal communication). The data presented here indicate that the dominant ITS copy in each dinoflagellate species could serve as an additional locus to COI or COB genes for DNA barcoding of dinoflagellate species. The advantage of using ITS region sequences is that this region has been sequenced extensively in dinoflagellates relative to COI.

Regardless of whether the ITS region, COI gene, or other sequences are selected, uniform criteria should be established to molecularly identify dinoflagellates and other algal species. We propose that ITS region sequences obtained from free-living dinoflagellates varying by $p > 0.04$ be used in helping to identify species. In genetically diverse assemblages this would allow identification of most species, including the potentially cryptic *Scrippsiella* species described by Montresor et al. (2003) (Fig. 3A). Inclusion of ITS or other identifying sequence data when strains or species are described in the literature, or deposited in culture collections, would facilitate the development of a functional dinoflagellate identification data base. This data base would allow researchers to rapidly and cost effectively check the taxonomic status of isolates before undertaking critical physiological, biochemical, or ecological studies. The ITS-based identification could also help ensure the long-term integrity of culture collections, make taxo-

nomic reassignments easier to track, and allow investigators to rapidly confirm whether a given isolate or individual cell type has been previously observed.

Support was provided by National Centers for Coastal Ocean Science, National Ocean Service, National Oceanic Atmospheric Administration, and the EPA ECOHAB STAR grant R826791-01-0 to K. Reece and N. Stokes. We would like to thank Todd LaJeunesse, Bill Sunda, and Patrice Mason for their constructive critiques of the manuscript. Debashish Bhattacharya also provided helpful suggestions on a previous draft of the paper. References to trade names do not imply product endorsement by the National Ocean Service, NOAA. ECOHAB contribution number 203, Virginia Institute of Marine Science contribution number 2787.

Adachi, M., Sako, Y. & Ishida, Y. 1996. Identification of the toxic dinoflagellates *Alexandrium catenella* and *A. tamarense* (Dinophyceae) using DNA probes and whole-cell hybridization. *J. Phycol.* 32:1049–52.

Adachi, M., Sako, Y. & Ishida, Y. 1997. Analysis of *Gymnodinium catenatum* dinophyceae using sequences of the 5.8S rDNA-ITS regions and random amplified polymorphic DNA. *Fish. Sci. (Tokyo)* 63:701–7.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E. & Tollervey, D. 1999. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO (Eur. Mol. Biol. Organ.) J.* 18:5399–410.

Allmang, C., Mitchell, P., Petfalski, E. & Tollervey, D. 2000. Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res.* 28:1684–91.

Biological Conservation Newsletter. 2004. Barcode of life initiative comes to life. http://ravenel.si.edu/bcn/issue/233.cfm#top. Accessed December 19, 2006. p. 1.

Cariello, N. F., Swenberg, J. A. & Skopek, T. R. 1991. Fidelity of *Thermococcus litoralis* DNA polymerase (Vent™) in PCR determined by denaturing gel electrophoresis. *Nucleic Acids Res.* 19:4193–8.

Cho, E. S., Kim, G.-Y. & Cho, Y. C. 2001. Molecular analysis of morphologically similar dinoflagellates *Cochlodinium polykrikoides*, *Gyrodinium impudicum* and *Gymnodinium catenatum* based on internal transcribed spacer and 5.8S rDNA regions. *Algae* 16:53–7.

D'Onofrio, G., Marino, D., Bianco, L., Busico, E. & Montresor, M. 1999. Toward an assessment on the taxonomy of dinoflagellates that produce calcareous cysts (Calciodinelloideae, Dinophyceae): a morphological and molecular approach. *J. Phycol.* 35:1063–78.

Daugbjerg, N., Hansen, G., Larsen, J. & Moestrup, O. 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* 39:302–17.

Edvardsen, B., Shalchian-Tabrizi, K., Jakobsen, K. S., Medlin, L. K., Dahl, E., Brubak, S. & Paasche, E. 2003. Genetic variability and molecular phylogeny of *Dinophysis* species (Dinophyceae) from Norwegian waters inferred from single cell analyses of rDNA. *J. Phycol.* 39:395–408.

Fensome, R. A., Saldarriaga, J. F. & Taylor, F. J. R. M. 1999. Dinoflagellate phylogeny revisited: reconciling morphological and molecular based phylogenies. *Grana* 38:66–80.

Galluzzi, L., Penna, A., Bertozzini, E., Vila, M., Garces, E. & Magnani, M. 2004. Development of a real-time PCR assay for rapid detection and quantification of *Alexandrium minutum* (a dinoflagellate). *Appl. Environ. Microbiol.* 70:1199–206.

Gerbi, S. A. 1986. The evolution of eukaryotic ribosomal DNA. *Biosystems* 19:247–58.

Giacobbe, M. G., Penna, A., Ceredi, A., Milandri, A., Poletti, R. & Yang, X. 2000. Toxicity and ribosomal DNA of the dinoflagellate *Dinophysis sacculus* (Dinophyta). *Phycologia* 39:177–82.

Giacobbe, M. G. & Yang, X. 1999. The life history of *Alexandrium taylori* (Dinophyceae). *J. Phycol.* 35:331–8.

Godhe, A., Anderson, D. M. & Rehnstam-Holm, A.-S. 2002. PCR amplification of microalgal DNA for sequencing and species identification: studies on fixatives and algal growth stages. *Harmful Algae* 1:375–82.

Haley, S. T., Cavender, J. F. & Murray, T. E. 1999. Detection of *Alexandrium tamarensis* by rapid PCR analysis. *Biotechniques* 26:88–91.

Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B Biol. Sci.* 270:S96–S9.

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. 2004. Identification of birds through DNA barcodes. *PLoS Biol.* 2:1657–63.

Hogg, I. D. & Hebert, P. D. N. 2004. Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can. J. Zool.* 82: 749–54.

Hudson, D. A. & Adlard, R. D. 1996. Nucleotide sequence determination of the partial SSU rDNA gene and ITS1 region of *Hematodinium* cf. *perezi* and *Hematodinium*-like dinoflagellates. *Dis. Aquat. Org.* 24:55–60.

Jorgensen, M. F., Murray, S. & Daugbjerg, N. 2004. *Amphidinium* revisited. I. Redefinition of *Amphidinium* (Dinophyceae) based on cladistic and molecular phylogenetic analyses. *J. Phycol.* 40:351–65.

Kovarik, A., Matyasek, R., Lim, K. Y., Skalicka, K., Koukalova, B., Knapp, S., Chase, M. & Leitch, A. R. 2004. Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana allotetraploids*. *Biol. J. Linn. Soc.* 82:615–25.

LaJeunesse, T. C., Loh, W. K. W., van Woesik, R., Hoegh-Guldberg, O., Schmidt, G. W. & Fitt, W. K. 2003. Low symbiont diversity in southern Great Barrier Reef corals, relative to those of the Caribbean. *Limnol. Oceanogr.* 48:2046–54.

LaJeunesse, T. C., Thornhill, D. J., Cox, E. F., Stanton, F. G., Fitt, W. K. & Schmidt, G. W. 2004. High diversity and host specificity observed among symbiotic dinoflagellates in reef coral communities from Hawaii. *Coral Reefs* 23:596–603.

LaJeunesse, T. C. & Trench, R. K. 2000. Biogeography of two species of *Symbiodinium* (Freudenthal) inhabiting the intertidal sea anemone *Anthopleura elegantissima* (Brandt). *Biol. Bull.* 199: 126–34.

Larsson, E. & Jacobsson, S. 2004. Controversy over *Hygrophorus cossus* settled using ITS sequence data from 200 year-old type material. *Mycol. Res.* 108:781–6.

Le Blancq, S. M., Khramtsov, N. V., Zamani, F., Upton, S. J. & Wu, T. W. 1997. Ribosomal RNA gene organization in *Cryptosporidium parvum*. *Mol. Biochem. Parasitol.* 90:463–78.

Lee, H.-O., Choi, K.-H. & Han, M.-S. 2003. Spring bloom of *Alexandrium tamarense* in Chinhae Bay, Korea. *Aquat. Microb. Ecol.* 33:271–8.

Litaker, R. W., Tester, P. A., Colorni, A., Levy, M. G. & Noga, E. J. 1999. The phylogenetic relationship of *Pfiesteria piscicida*, Cryptoperidiniopsoid sp., *Amyloodinium ocellatum* and a *Pfiesteria*-like dinoflagellate to other dinoflagellates and apicomplexans. *J. Phycol.* 35:1379–89.

Litaker, R. W., Vandersea, M. W., Kibler, S. R., Reece, K. S., Stokes, N. A., Steidinger, K. A., Millie, D. F., Bendis, B. J., Pigg, R. J. & Tester, P. A. 2003. Identification of *Pfiesteria piscicida* (Dinophyceae) and *Pfiesteria*-like organisms using internal transcribed spacer-specific PCR assays. *J. Phycol.* 39:754–61.

Lygerou, Z., Allmang, C., Tollervey, D. & Seraphin, B. 1996. Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro. *Science* 272:268–70.

Marin, I., Aguilera, A., Reguera, B. & Abad, J. P. 2001. Preparation of DNA suitable for PCR amplification from fresh or fixed single dinoflagellate cells. *Biotechniques* 30:88–93.

Maroteaux, L., Herzog, M. & Soyer-Gobillard, M. O. 1985. Molecular organization of dinoflagellate ribosomal DNA: evolutionary implications of the deduced 5.8s ribosomal RNA secondary structure. *Biosystems* 18:307–20.

Montresor, M., Sgrosso, S., Procaccini, G. & Kooistra, W. H. C. F. 2003. Intraspecific diversity in *Scrippsiella trochoidea* (Dinophyceae): evidence for cryptic species. *Phycologia* 42:56–70.

Moreira, D. & Lopez-Garcia, P. 2002. The molecular ecology of microbial eukaryotes unveils a hidden world. *Trends Microbiol.* 10:31–8.

Nagai, S., Itakura, S., Matsuyama, Y. & Kotani, Y. 2003. Encystment under laboratory conditions of the toxic dinoflagellate *Alexandrium tamiyavanichii* (Dinophyceae) isolated from the Seto Inland Sea, Japan. *Phycologia* 42:646–53.

Nazar, R. N. 1984. The ribosomal 5.8S RNA: eukaryotic adaptation or processing variant? *Can. J. Biochem. Cell Biol.* 62:311–20.

Nazar, R. N. 2004. Ribosomal RNA processing and ribosome biogenesis in eukaryotes. *IUBMB Life* 56:457–65.

Ohta, T. 2000. Evolution of gene families. *Gene* 259:45–52.

Ordas, M. C., Fraga, S., Franco, J. M., Ordas, A. & Figueras, A. 2004. Toxin and molecular analysis of *Gymnodinium catenatum* (Dinophyceae) strains from Galicia (NW Spain) and Andalucia (S Spain). *J. Plankton Res.* 26:341–9.

Penna, A. & Magnani, M. 1999. Identification of *Alexandrium* (Dinophyceae) species using PCR and rDNA-targeted probes. *J. Phycol.* 35:615–21.

Pin, L. C., Teen, L. P., Ahmad, A. & Usup, G. 2001. Genetic diversity of *Ostreopsis ovata* (Dinophyceae) from Malaysia. *Mar. Biotechnol.* 3:246–55.

Probert, I., Lewis, J. & Erard-Le Denn, E. 2002. Morphological details of the life history of *Alexandrium minutum* (Dinophyceae). *Cryptogam. Algol.* 23:343–55.

Rappe, M. S., Vergin, K. & Giovannoni, S. J. 2000. Phylogenetic comparisons of a coastal bacterioplankton community with its counterparts in open ocean and freshwater systems. *FEMS Microbiol. Ecol.* 33:219–32.

Rooney, A. P. 2004. Mechanisms underlying the evolution and maintenance of functionally heterogeneous 18S rRNA genes in apicomplexans. *Mol. Biol. Evol.* 21:1704–11.

Rosenberg, M. S. 2005. Multiple sequence alignment accuracy and evolutionary distance estimation. *BMC Bioinformatics* 6:278.

Sako, Y. 1999. Molecular identification of toxic marine dinoflagellates *Alexandrium* species. *Arch. Hydrobiol.* 54:311–24.

Saldarriaga, J. F., Taylor, F., Cavalier-Smith, T., Menden-Deuer, S. & Keeling, P. J. 2004. Molecular data and the evolutionary history of dinoflagellates. *Eur. J. Protistol.* 40:85–111.

Santos, S. R., Kinzie, R. A. III, Sakai, K. & Coffroth, M. A. 2003. Molecular characterization of nuclear small subunit (18S)-rDNA pseudogenes in a symbiotic dinoflagellate (*Symbiodinium*, Dinophyta). *J. Eukaryot. Microbiol.* 50:417–21.

Scholin, C. A. & Anderson, D. M. 1994. Identification of group- and strain-specific genetic markers for globally distributed *Alexandrium* (Dinophyceae) I. RFLP analysis of SSU rRNA genes. *J. Phycol.* 30:744–54.

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). *J. Phycol.* 29:209–16.

Scholin, C. A., Hallegraeff, G. M. & Anderson, D. M. 1995. Molecular evolution of the *Alexandrium tamarense* 'species complex' (Dinophyceae): dispersal in the North American and West Pacific regions. *Phycologia* 34:472–85.

Shankle, A. M., Mayali, X. & Franks, P. J. S. 2004. Temporal patterns in population genetic diversity of *Prorocentrum micans* (Dinophyceae). *J. Phycol.* 40:239–47.

Silva, E. S. & Faust, M. A. 1995. Small cells in the life history of dinoflagellates (Dinophyceae): a review. *Phycologia* 34:396–408.

Sournia, A. 1986. *Atlas du phytoplancton marin. Volume 1: Cyanophycées, Dictyochophycées, Dinophycées, Raphidophycées*. Centre National de la Recherche Scientifique, Paris, 219 pp.

Steidinger, K., Landsberg, J. H., Richardson, R. W., Truby, E., Blakesley, B., Scott, P., Tester, P. A., Tengs, T., Mason, P., Morton, S. L., Seaborn, D., Litaker, W., Reece, K., Oldach, D., Haas, L. & Vasta, G. R. 2001. Classification and identification of *Pfiesteria* and *Pfiesteria*-like species. *Environ. Health Perspect.* 109 (Suppl. 5): 661–5.

Steidinger, K. A., Mason, P. L., Vogelbein, W. K., Tester, P. A. & Litaker, R. W. 2006. *Cryptoperidiniopsis brodyi* (Dinophyceae), a small lightly armored dinoflagellate in the Pfiesteriaceae. *J. Phycol.* 42:951–61.

Suzuki, M. T., Beja, O., Taylor, L. T. & DeLong, E. F. 2001. Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton. *Environ. Microbiol.* 3:323–31.

Swofford, D. L. 2002. *PAUP*: Phylogenetic Analysis using Parsimony (*and Other Methods), Version 4b10*. Sinauer Associates, Sunderland, MA, USA.

Takabayashi, M., Santos, S. R. & Cook, C. B. 2004. Mitochondrial DNA phylogeny of the symbiotic dinoflagellates (*Symbiodinium*, Dinophyta). *J. Phycol.* 40:160–4.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 15:4876–82.

Tindall, K. R. & Kunkel, T. A. 1988. Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* 27: 6008–13.

van Oppen, M. J. H., Palstra, F. P., Piquet, A. M. T. & Miller, D. J. 2001. Patterns of coral-dinoflagellate associations in Acropora: significance of local availability and physiology of *Symbiodinium* strains and host-symbiont selectivity. *Proc. R. Soc. Lond. B Biol. Sci.* 268:1759–67.

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., 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., Pfannkoch, C., Rogers, Y. H. & Smith, H. O. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304:66–74.

Wagener, B., Reineke, A., Loehr, B. & Zebitz, C. P. W. 2004. A PCR-based approach to distinguish important *Diadegma* species (Hymenoptera: Ichneumonidae) associated with diamondback moth, taxonomic *Plutella xylostella* (Lepidoptera: Plutellidae). *Bull. Entomol. Res.* 94:465–71.

Wagstaff, S. J. 2004. Evolution and biogeography of the austral genus *Phyllocladus* (Podocarpaceae). *J. Biogeogr.* 31:1569–77.

Zhang, Z., Green, B. R. & Cavalier-Smith, T. 2000. Phylogeny of ultra-rapidly evolving dinoflagellate chloroplast genes: a possible common origin for sporozoan and dinoflagellate plastids. *J. Mol. Evol.* 51:26–40.

Zhang, H. & Lin, S. J. 2002. Detection and quantification of *Pfiesteria piscicida* by using the mitochondrial cytochrome b gene. *Appl. Environ. Microbiol.* 68:989–94.

Zingone, A., Montresor, M. & Marino, D. 1998. Morphological variability of the potentially toxic dinoflagellate *Dinophysis sacculus* (Dinophyceae) and its taxonomic relationships with *D. pavillardii* and *D. acuminata*. *Eur. J. Phycol.* 33:259–73.

Supplementary Material

The following supplementary material is available for this article:

Appendix S1. GenBank numbers for sequences used in the analysis.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1529-8817.2007.00320.x

(This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.