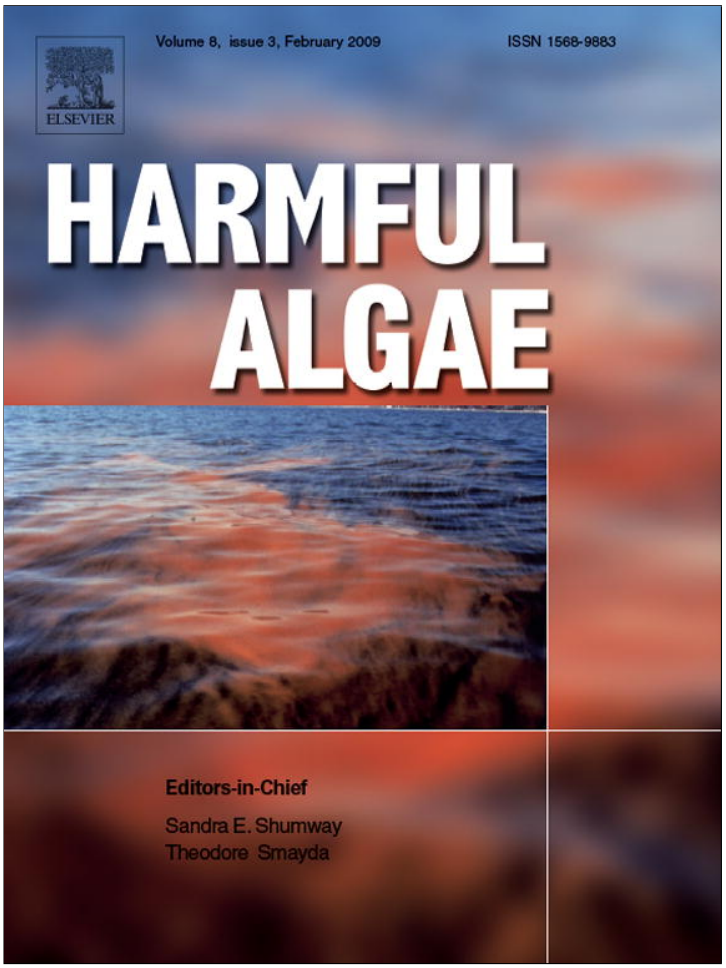

Harmful Algae 8 (2009) 441–446

Contents lists available at ScienceDirect

Harmful Algae

journal homepage: www.elsevier.com/locate/hal

Secondary structural modeling of the second internal transcribed spacer (ITS2) from *Pfiesteria*-like dinoflagellates (Dinophyceae)

Kevin M. Kocot, Scott R. Santos *

101 Rouse Life Sciences Building, Department of Biological Sciences and Cellular and Molecular Biosciences Peak Program, Auburn University, Auburn, AL 36849, USA

ARTICLE INFO

Article history:
Received 12 February 2008
Received in revised form 15 September 2008
Accepted 16 September 2008

Keywords:
Cryptoperidiniopsis
Dinoflagellate
ITS2
Pfiesteria
rRNA
Secondary structure

ABSTRACT

Pfiesteria piscicida is a harmful bloom-forming alga that has received a great deal of attention due to its potential association with large fish kills and neurological problems in humans. Since the discovery of *Pfiesteria*, several other *Pfiesteria*-like dinoflagellates (PLDs) have also been identified. Genetic identification and phylogenetic relationships among the PLDs commonly utilize sequence data from the genes and spacers of the ribosomal DNA (rDNA) operon. Of these, the internal transcribed spacers (ITSs) have been previously shown to fold into secondary structures that are critical for proper ribosomal processing. In this study, we modeled the secondary structure of the second internal transcribed spacer (ITS2) from 16 PLDs (as well as an outgroup taxon) using phylogenetic comparative methods and minimum free energy. The secondary structural models predicted for these dinoflagellates consisted of four paired helices separated by five unpaired regions, consistent with those reported from many eukaryotes. All of the structures were highly stable ($\Delta G = -66.1$ to -122.3 kcal·mol at 37 °C) and several structural characters were found to be conserved either across the PLDs or were specific to monophyletic subgroups, strengthening previously inferred phylogenetic relationships among taxa. Additionally, an 18 bp motif was identified in the PLDs whose position corresponds to a ribosomal processing site described from other eukaryotes. Potential applications of these ITS2 secondary structures include utility in strain and species identification, phylogenetic inference and serving as a tool for identifying and excluding rDNA pseudogenes when assessing biodiversity within the PLDs.

1. Introduction

Dinoflagellates are a diverse and important group of aquatic microalgae. One species, *Pfiesteria piscicida* Steidinger and Burkholder, has received a great deal of attention due to its potential association with large fish kills and neurotoxicosis in humans (Burkholder and Glasgow, 1997; Marshall et al., 2000; Moeller et al., 2001; Gordon et al., 2003; Burkholder et al., 2005; Gordon and Dyer, 2005; Moeller et al., 2007). Because of the ecological and economic losses associated with toxins from at least some strains of *Pfiesteria*, extensive environmental sampling has been conducted toward understanding the distribution and diversity of this harmful alga. As a result, other *Pfiesteria*-like dinoflagellates have been discovered and described on the basis of both morphological and molecular characters. For example, recently described relatives of *Pfiesteria* include *Stockeria algicida* (Jeong et al., 2005), *Cryptoperidiniopsis brodyi* (Steidinger et al., 2006), *Luciella masanensis*, and *Luciella atlantis* (Mason et al., 2007). Additionally, several other *Pfiesteria*-like dinoflagellates are known, but remain to be formally described (e.g., Burkholder and Glasgow, 1997; Litaker et al., 1999; Seaborn et al., 2006). Phylogenetic studies of *Pfiesteria*, *Cryptoperidiniopsis*, *Luciella* and other (undescribed) *Pfiesteria*-like dinoflagellates indicate that they form a monophyletic group within the order Peridiniales and are closely related to calcareous dinoflagellates such as *Thoracosphaera* (see Marshall et al., 2006 for discussion). Here, we refer to *Pfiesteria*, *Cryptoperidiniopsis*, *Luciella*, and other undescribed relatives of *Pfiesteria* collectively as "*Pfiesteria*-like dinoflagellates," or PLDs, for short.

The genetic identification and phylogenetics of PLDs commonly utilize the genes and spacers of the ribosomal RNA (rRNA) operon. In eukaryotes, this operon occurs as tandem repeats in the nuclear genome, each possessing three coding regions (i.e., the 18S-, 5.8S-, and 28S-rDNAs) and three non-coding spacers (Hillis and Dixon, 1991). These spacers, the external transcribed spacer (ETS) and the internal transcribed spacers 1 (ITS1) and 2 (ITS2), are transcribed with the coding regions as a single precursor rRNA (pre-rRNA) molecule. At higher taxonomic levels, the primary DNA sequence and length of the ITSs are typically highly variable (Coleman, 2003,

* Corresponding author. Tel.: +1 334 844 7410; fax: +1 334 844 1645.
E-mail address: santos@auburn.edu (S.R. Santos).

1568-9883/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.hal.2008.09.004

2007). However, these characteristics have proven useful for species differentiation (e.g., Furlong and Maden, 1983; Connell, 2001; Lu et al., 2002; Luo and Mitchell, 2002; Noge et al., 2005) as well as for inferring phylogenetic relationships (e.g., Baldwin, 1992; Mai and Coleman, 1997; Coleman and Mai, 1997; Weekers et al., 2001; Coleman and Vacquier, 2002) in many groups of closely related taxa.

Although their primary DNA sequence and length are poorly conserved across broad taxonomic groups, the secondary structures of the ITSs possess features that are ubiquitous in virtually all eukaryotes (Coleman, 2003; Schultz et al., 2005; Coleman, 2007). For instance, secondary structural analyses of ITS2 from diverse eukaryotes (e.g., Michot et al., 1999; Mai and Coleman, 1997; Joseph et al., 1999; Goertzen et al., 2003; Chen et al., 2004) have revealed that the molecule generally folds into four double-stranded regions (termed helices or hairpins) separated by single-stranded regions. The importance of this secondary structure has been demonstrated in *Drosophila*, rodents, and yeast (*Saccharomyces cerevisiae*) (van der Sande et al., 1992; Schlötterer et al., 1994; van Nues et al., 1995; Good et al., 1997); in the case of yeast, nucleotide deletions that disrupt ITS2 secondary structure significantly decrease the production of ribosomal subunits and overall growth rates (van Nues et al., 1995). Thus, although ITS2 has previously been considered "junk DNA," its structural characteristics are a vital component in the maturation of the rRNA subunits and overall function of the cell.

Previously, Gottschling and Plötner (2004) presented secondary structural models for the ITS regions of several dinoflagellate taxa, including *P. piscicida*. Overall, the ITS2 secondary structure of the examined dinoflagellates was similar to that described from other eukaryotes. However, the presence of an additional helix, termed subhelix IIIa, was identified in a member of the symbiotic dinoflagellate genus *Symbiodinium* (Gottschling and Plötner, 2004). To investigate whether this structural difference was consistent across *Symbiodinium*, Hunter et al. (2007) generated ITS2 models for all eight clades of these dinoflagellates. It was found that ITS2 in *Symbiodinium* clades A, D, E and G folds into the general four-helix conformation while the five-helix conformation reported by Gottschling and Plötner (2004) is an apomorphy of a monophyletic group consisting of clades B, C, F, and H. Other conserved features, including a processing site corresponding to one previously known from other taxa (see Wesson, 1992 for discussion), were also identified (Hunter et al., 2007). Thus, rRNA secondary structural models can prove useful for understanding the molecular evolution of this region, as an aid in the design of nucleic acid probes for *in situ* hybridization, and/or as a tool for identifying and excluding rDNA pseudogenes in diversity assessments (Thornhill et al., 2007). While such analyses could also prove useful in studies of PLDs, ITS2 structural models are a prerequisite for such work. In this study, we modeled the structural features of the ITS2 region from 16 PLDs isolates as well as an outgroup species, *Karlodinium micrum* Leadbeater and Dodge. Along with furthering knowledge on the biology of these dinoflagellates, we feel that the ITS2 secondary structures presented here have utility in strain and species identification, estimates of biodiversity, and strengthen previously inferred phylogenetic relationships among the PLDs.

2. Materials and methods

2.1. ITS2 sequences from the PLDs

The ITS2 sequences used in this study were acquired from GenBank and are listed in Table 1. These sequences were generated either by direct sequencing of PCR products obtained from clonal cultures (Litaker et al., 2003, 2005; Litaker et al. unpublished data; Steidinger et al., 2006; Mason et al., 2007) or via bacterial cloning (Litaker et al., 2005). Average GC content for each sequence and pairwise distances between all sequences were calculated using PAUP* v4.0 b10 (Swofford, 2002).

2.2. Determination of ITS2 secondary structure from the PLDs

Structural determination for the selected sequences was initially done by comparing each to the ITS2 secondary structure of *P. pisicida* as predicted by Gottschling and Plötner (2004), utilizing the phylogenetic comparative method described by Mai and Coleman (1997). Specifically, sequences were first aligned to the ITS2 sequence of *P. pisicida* utilized by Gottschling and Plötner (2004). Spans of nucleotides from the examined sequences were then selected, using the *P. pisicida* ITS2 structure as a guide, and folded using MFOLD v3.2 (Mathews et al., 1999; Zuker, 2003) under the default settings. Resulting helices were compared to each other to identify common structural characteristics and/or potentially misfolded regions. Once the conformation of each individual helix was determined, the entire ITS2 secondary

Table 1
Secondary structural characteristics of the proposed *Pfiesteria*-like dinoflagellate (PLDs) ITS2 models.

Species, isolate, culture #, and/or ribotype	GenBank accession #	ITS2 length	*a*	I[a]	*b*	II[a]	*c*	III[a]	*d*	IV[a]	*e*	ΔG[b]	GC (%)
Pfiesteria piscicida	AF352333	244	10	21 + 4(3)	7	16 + 3(2)	10	30 + 4(6)	15	9 + 4(0)	4	−97.6	63.11
Pfiesteria shumwayae VIMS 1049	AY245694	324	11	18 + 4(2)	7	43 + 3(8)	9	31 + 5(6)	20	8 + 5(0)	4	−122.3	58.02
Pfiesteria-like sp. F525Jul02	AY590480	237	11	19 + 5(2)	7	20 + 4(3)	9	30 + 4(4)	15	5 + 3(0)	4	−86.8	49.37
Pfiesteria-like sp. HR1NovC5	AY590482	249	11	21 + 6(4)	7	17 + 4(2)	9	30 + 6(5)	13	10 + 4(1)	3	−97.1	54.62
Luciella masanensis ribotype 1	AY590485	305	11	21 + 7(3)	8	18 + 4(2)	3	36 + 6(7)	9	24 + 4(4)	6	−115.3	53.11
Luciella masanensis ribotype 2	AY245689	245	11	19 + 4(4)	7	18 + 4(2)	9	31 + 4(6)	13	9 + 4(1)	3	−88.0	53.06
Luciella masanensis ribotype 4	AY590483	246	11	19 + 4(2)	7	19 + 4(2)	10	28 + 4(4)	17	8 + 4(1)	4	−86.9	55.28
Luciella masanensis ribotype 3	AY590477	265	11	21 + 4(3)	7	23 + 5(2)	9	30 + 4(5)	18	8 + 4(1)	3	−90.6	48.62
Luciella atlantis	AY590478	259	11	21 + 4(4)	7	22 + 2(5)	9	29 + 4(4)	18	9 + 4(1)	3	−109.5	62.16
Cryptoperidiniopsis brodyi CCMP 1828	AY590476	248	11	21 + 3(3)	7	21 + 5(2)	9	30 + 4(5)	14	5 + 5(0)	4	−90.3	50.00
Cryptoperidiniopsis brodyi V14	AY245690	297	11	20 + 5(3)	7	21 + 4(2)	12	26 + 6(4)	24	25 + 4(4)	4	−106.7	52.19
Cryptoperidiniopsis sp. Folly C5	AY590481	257	11	19 + 3(3)	9	20 + 4(3)	11	30 + 6(6)	8	10 + 6(1)	4	−87.1	53.70
Cryptoperidiniopsis sp. A5	AF352355	252	11	19 + 5(4)	7	19 + 4(1)	9	31 + 4(6)	20	6 + 6(0)	4	−89.2	53.97
Cryptoperidiniopsis sp. NOAA Beach	AY590486	309	11	20 + 5(4)	7	34 + 4(6)	11	26 + 4(5)	25	6 + 6(0)	4	−116.9	60.19
"Shepard's Crook"	AY590479	197	11	14 + 4(2)	7	14 + 3(1)	9	30 + 3(5)	7	5 + 3(0)	5	−66.1	44.67
"Shepard's Crook"	AY590484	197	8	16 + 4(2)	2	16 + 5(2)	1	29 + 3(6)	0	7 + 4(1)	4	−73.3	46.70
Karlodinium micrum	AF352367	218	8	21 + 4(3)	3	14 + 6(1)	8	24 + 4(5)	9	7 + 6(0)	5	−78.1	49.54

[a] $X + Y(Z)$ where X is the length of the helix in paired bases, Y is the number of unpaired bases at the apex (i.e., loop) of the helix, and Z is the number of unpaired regions in the helix.
[b] ΔG = kcal·mol at 37 °C.

structure of that isolate was assembled using RNAstructure v4.5 (Mathews et al., 2004) and visualized using loopDloop v1.2a64 (Gilbert, 1992). The free energy (ΔG = kcal·mol at 37 °C) of each structure was calculated using RNAstructure v4.5.

3. Results

3.1. General primary sequence characteristics of ITS2 in the PLDs

The examined ITS2 sequences varied in length from 197 to 324 nucleotides (nt) (Litaker et al., 2003, 2005; Litaker et al. unpublished data; Steidinger et al., 2006; Mason et al., 2007; Table 1). Our start and end positions of ITS2 were chosen based on the sequence alignment and structural model of *P. piscicida* ITS2 proposed by Gottschling and Plötner (2004). Thus, the sizes reported for some sequences in this study are slightly different from their annotations in GenBank. It should be noted that this has no effect on the predicted secondary structures as the endpoints occur in unpaired regions (i.e., *a* and *e*, explained below). Average GC content and pairwise distances of the 17 sequences used here were 53% and 0.39 uncorrected (i.e., *p*) distances, respectively. Exclusion of the outgroup taxon, *K. micrum*, only slightly reduced the average pairwise distance to 0.37 *p*, suggesting this outgroup is not highly divergent from the PLDs examined here. These PLD ITS2 lengths and GC contents fall within the range of values previously reported for other dinoflagellates (Gottschling and Plötner, 2004; Hunter et al., 2007; Thornhill et al., 2007).

3.2. Secondary structural characteristics of ITS2 in the PLDs

Following the nomenclature of Gottschling and Plötner (2004), the secondary structural models predicted in this study possess four paired regions, or helices (designated I–IV), separated by single-stranded regions (designated *a*–*e*) (Fig. 1). Only these four helices were recovered in all cases. Each terminated in an apical loop of three or more unpaired nt while I–III were characterized by at least one unpaired region, such as a mismatch (two opposing unpaired nucleotides), an internal loop (multiple opposing unpaired nucleotides), or a bulge (one or more adjacent unpaired nucleotides) in a helix.

Although the general secondary structure of ITS2 was well conserved across the 17 taxa, variation was observed between models. For instance, helix lengths (in paired bases, bp) ranged in size from 14 to 43 bp for helix II to 5–25 bp for helix IV (Table 1). In general, helix III was the longest while helix IV was the shortest (Table 1). Exceptions to this were noted for some isolates, where

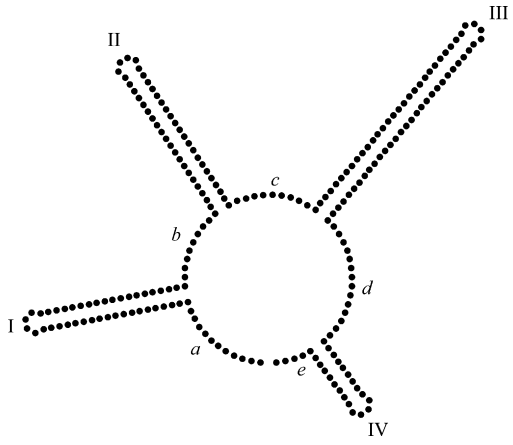

Fig. 1. Consensus secondary structure diagram for ITS2 of the PLDs used in this study.

apparent insertions increased the overall length of that helix. Examples include helix II of *Pfiesteria shumwayae* (AY245694) and *Cryptoperidiniopsis* sp. ribotype NOAA Beach (AY590486), which were 43 and 34 bp, respectively. Length exceptions where also noted for helix IV of *C. brodyi* ribotype V14 (AY245690) and *L. masanensis* ribotype 1 (AY590485) (25 and 24 bp, respectively; Table 1). Moreover, the number of unpaired regions (excluding apical loops) varied within each helix (Table 1). The inter-helix core regions were also of varying lengths (in unpaired nucleotides, nt; region *a*: 8–11 nt; region *b*: 2–9 nt; region *c*: 1–12 nt; region *d*: 0–25 nt; region *e*: 3–6 nt), and in some cases, completely or nearly absent (e.g., "Shepard's Crook" (AY590484); Table 1). In spite of this variation, all of the proposed ITS2 secondary structures were highly stable. Free energy values (ΔG) varied between −66.1 and −122.3 kcal·mol at 37 °C (Table 1), similar to that of other dinoflagellates (Gottschling and Plötner, 2004; Hunter et al., 2007).

Additionally, several conserved features were observed across the ITS2 secondary structures. One such feature was a pyrimidine–pyrimidine bulge in helix II (boxed in Supplementary Fig. S1), which is considered as a hallmark of this helix since it is found in most ITS2 structures (Coleman and Vacquier, 2002; Coleman, 2003, 2007). Pyrimidine-rich loop(s) were also observed within helix III of all taxa. For most, these loops contained >50% pyrimidine residues and were larger on the 5′ (relative to the 3′) side of the helix (Supplementary Fig. S1). This feature is also apparent in the ITS2 secondary structures predicted for representatives of the Calciodinelloideae, Thoracosphaeraceae, Heterocapsaceae and *Symbiodinium* (Gottschling and Plötner, 2004; Hunter et al., 2007). Another identified feature of nearly all the PLD ITS2 secondary structures was the presence of two adjacent unpaired nucleotides near the 3′ base of helix I. These nucleotides were always pyrimidines and predominantly uracil residues; however, three of five *Luciella* isolates possessed a cytosine at one of these positions. Notably, this feature was absent from the two "Shepard's Crook" isolates and *K. micrum* (O–Q of Supplementary Fig. S1). Along with this, a loop near the base of helix IV was found in the *Luciella* isolates, with the exception being *L. masanensis* ribotype 1, which differs due to a nucleotide insertion producing additional loops in that helix (Supplementary Fig. S1C). A region of high nucleotide conservation, in the form of a nearly identical 18 nt motif on the 3′ side of helix II spanning alignment positions 182–199 occurred in all 17 models. The consensus sequence of this motif is 5′-GCCUUUGACACAUUNAAG-3′ and is 100% conserved at seven positions, >80% conserved at six positions, and >60% conserved at four positions (Fig. 2) across these PLDs.

Some of the sequences examined here contained nucleotide ambiguities at particular positions. Specifically, ambiguities were observed in *C. brodyi* ribotype V14 (AY245690), *P. shumwayae* (AY245694), *L. atlantis* (AY590478), *L. masanensis* ribotype 1 (AY590485), *L. masanensis* ribotype 2 (AY245689), and *L. masanensis* ribotype 3 (AY590477), and are shown in red in Supplementary Fig. S1. We found that 15 of the 31 ambiguities occurred in unpaired regions of the models, thus having no impact on the predicted secondary structures. On the other hand, nine ambiguities had the potential to alter structural features. However, eight of these occurred in either single-stranded regions within helices or directly adjacent to the base of a helix, which would change the size of a loop or length of a helix by only 1 bp, respectively. Lastly, seven ambiguities were identified as hemi-compensatory base changes (CBCs), single base changes that maintain stable hydrogen bonded pairs (e.g., G–C vs. G–U) within a proposed helix.

In addition to the sequences used to construct the secondary structural models presented here (Supplementary Fig. S1), multiple ITS2 sequences were available on GenBank for six of the taxa

GenBank no.	Motif
AF352333	GCCUUUGACGCAUUGAGG
AY245694	GCCUUUGACGCAUUGAGG
AY590480	GCCUUUGACGCAUUUAAA
AY590482	GCCCUUGACACAUCUAGG
AY590485	GCCUUUGACACAYCUAAG
AY245689	GCCUUUGACACAUCUAAG
AY590483	GCCUUUGACACAUCCAAG
AY590477	GCCUUUGACACAUCUAAG
AY590478	GCCCUUGACGCGUUGAAG
AY590476	GCCUUUGACACAUUGAAG
AY245690	CCUUUUGACRCGUUUAAG
AY590481	GCCUUUGACACGUUAAAG
AF352355	UUCUUUGACACGUUCAAG
AY590486	CCCUUUGACGCGCUUAAG
AY590479	GCCUUUGACACAUUGAAU
AY590484	GCCUUUGACACAUGAAGG
AF352367	GCCUUUGACGCAUUCAGU
	********** *** ********** * * * ***** * * *
Consensus	GCCUUUGACACAUUNAAG

Fig. 2. Alignment of the conserved 18 bp motif identified in helix II of the ITS2 across the PLDs. *, ≥60% conserved; **, ≥80% conserved; *** = 100% conserved.

examined in this study: *Cryptoperidiniopsis* sp. ribotype A5 (AF352356, AF352357, AF352358, AY245691), *C. brodyi* ribotype V14 (AF352349, AF352350, AF352351), *K. micrum* (AF352365, AF352366, AY245692), *L. masanensis* ribotype 2 (AF352346, AF352347, AF352348), *P. piscicida* (AF352334, AF352335, AF352336, AF352337, AY245693), and *P. shumwayae* (AF352338, AF352339, AF352340, AF352341, AF352342, AF352343, AF352344, AF352345). For each of these taxa, sequences were aligned and differences among them examined to determine their effect on ITS2 secondary structure and free energy values. Nucleotide positions were found to vary among the sequences from a given species or ribotype (a total of 12 positions across the six taxa). Of these variable sites, seven occurred in unpaired regions and did not affect the ITS2 secondary structure or calculated free energy. Four sites represented hemi-CBCs that also did not affect the proposed secondary structural models. Lastly, one nucleotide difference among the *L. masanensis* ribotype 2 sequences was the presence of a G vs. an A in stem III, at position 187. This leads to the occurrence of a G–C bond for the three isolates in which a G is present, whereas in the single isolate in which an A occurs leads to the formation of a 1 bp loop. Since this polymorphism is confined to a single sequence, it may represent sequencing error, a pseudogene or naturally occurring, but rare, genetic variation in this *L. masanensis* ribotype.

4. Discussion

The ITS2 secondary structures predicted here for the PLDs consisted of four helices separated by single-stranded regions. Thus, our models are consistent with the notion that ITS2 secondary structure is generally conserved throughout Eukarya (see Section 1). Along with this, structural features shared by particular clades of PLDs were also identified. For example, nearly all PLDs possess two adjacent unpaired pyrimidines at the 3′ base of helix I. However, this feature was absent from both "Shepard's Crook" isolates (and the outgroup taxon *K. micrum*) examined here, similar to secondary structures predicted for the Calciodinelloideae, Thoracosphaeraceae, Heterocapsaceae and *Symbiodinium* (Gottschling and Plötner, 2004; Hunter et al., 2007). Interestingly, recent phylogenetic analyses suggest that the "Shepard's Crook" isolates and *S. algicida* (Jeong et al., 2005) form a sister clade to one including *Pfiesteria*, *Luciella*, and *Cryptoperidiniopsis* (Marshall et al., 2006; Mason et al., 2007). The lack of the above feature in the "Shepard's Crook" isolates supports their distinction from other PLDs. Likewise, loop(s) of two or more unpaired nucleotides in helix IV were noted in all *Luciella* sp., which is a structural characteristic rarely observed in dinoflagellates outside of this genus (Gottschling and Plötner, 2004; Hunter et al., 2007). On the other hand, no obvious structural features appear to unite the *Cryptoperidiniopsis* isolates. Gottschling and Plötner (2004) reported that helix IV was extremely elongated in the cryptoperidiniopsoid isolate they examined, suggesting it might be a feature of the genus. In contrast, we found this only to be the case for two of the five cryptoperidiniopsoids examined in this study—*Cryptoperidiniopsis* sp. ribotype Folly C5 (AY590481, Supplementary Fig. S1L) and *C. brodyi* ribotype V14 (AY245690, Supplementary Fig. S1K).

Other evolutionarily conserved features were identified from our structural analyses of PLD ITS2 including a pyrimidine–pyrimidine bulge in helix II and a conserved motif corresponding to a previously described processing site. The pyrimidine–pyrimidine bulge in helix II is a nearly universal feature of ITS2 secondary structures and is thought to be important in processing of the molecule due to its sequence and positional conservation (Mai and Coleman, 1997; Coleman, 2003). Notably, in PLD ITS2 structures in which this feature is absent, a single nucleotide bulge occurs in its place and may be a vestige of this feature. We also identified a region of high-nucleotide conservation (colored green in Supplementary Fig. S1) in the same region of all the predicted ITS2 secondary structures. This 18 nt conserved motif, located between 78 and 95 nt or 169 and 186 nt from the 5′ start of ITS2, is consistent with a ribosomal processing site previously identified in yeast (Veldman et al., 1981) and rats (Reddy et al., 1983) and implicated to have a similar function in *Xenopus* (Hall and Maden, 1980), sea urchin (Hidenach and Stafford, 1984), flies (Tautz et al., 1988; Wesson, 1992), and parasitic flatworms (Morgan and Blair, 1998). Thus, this conserved region likely serves as a processing site in the PLDs as well. A conserved motif of 11 nt on the 3′ side of helix II has also been previously identified in *Symbiodinium* (Hunter et al., 2007). Interestingly, the conserved motif in the PLDs shares little to no homology with that of *Symbiodinium* or any of the above-mentioned taxa. This further supports that positional homology, rather than characteristics of the primary sequence itself, are important for proper ITS processing and the biogenesis of rRNA subunits (Joseph et al., 1999; Schultz et al., 2005).

Some of the sequences that we utilized possessed nucleotide ambiguities, where >1 nt was identified as occurring in the same position during sequencing. Because the rDNA operon occurs as a series of tandem repeats (as many as 1200 estimated in dinoflagellates; Galluzzi et al., 2004), ambiguities such as these are typically due to intragenomic variability between copies of the operon. Although ambiguities may be problematic in other situations, these can prove useful in elucidating rRNA secondary structures via the identification of CBCs and hemi-CBCs. CBCs are corresponding mutational changes that occur at nucleotide positions pairing in a double-stranded region. In contrast, hemi-CBCs are single base changes that also continue to maintain stable hydrogen bonding between pairs. In both cases, these types of mutations do not disrupt the secondary structure and it has been argued that they provide support for a proposed structural model (Kimura, 1985; Gutell et al., 1994; Savill et al., 2001; Coleman, 2003). Of the sequence ambiguities we identified, the majority

(71%) had no effect on the secondary structure of the molecule since they occurred in unpaired regions or were hemi-CBCs in nature. For the few ambiguities that did alter secondary structural characteristics, all but one only influenced the size of a loop within a helix, or the total length of a helix, by a single base pair. Thus, we feel that the non-random distribution of these ambiguities provides additional support for the proposed ITS2 secondary structures of the PLDs.

In addition to nucleotide ambiguities or variation within the ITS2 sequences themselves, comparison of ITS2 sequences from the same species, strain and/or ribotype revealed apparent intraspecific variation among sequences. Of these, 11 of 12 (92%) had no effect on the predicted secondary structural models. In only one case did a difference affect secondary structure, that of helix III of *L. masanensis* ribotype 2, resulting in the loss a G–C bond and the formation of 1 bp loop. Taken together, we feel that the non-random distribution of nucleotide variation observed among sequences from the same species, strain and/or ribotype, along with the nucleotide ambiguities within a single sequence discussed previously, provide strong support for the ITS2 secondary structures proposed here.

Identification and systematics for dinoflagellates such as the PLDs have traditionally relied on SEM examination of thecal plates using the tabulation system developed by Kofoid (1909). This procedure, however, is laborious, expensive, and time-consuming (Steidinger et al., 1996; Truby, 1997). With the proliferation of molecular techniques, nucleic acid (i.e., DNA) sequencing of regions from the rDNA operon is now commonly utilized in this pursuit. For the PLDs, phylogenies inferred by both morphological and molecular data are generally well correlated (Litaker et al., 1999; Taylor, 1999). Along with this, DNA sequencing is relatively rapid and straightforward and has proven useful for distinguishing morphologically similar and co-occurring species and/or strains of PLDs (Litaker et al., 2003, 2007; Seaborn et al., 2006). Unfortunately, given the multi-copy nature of the eukaryotic rDNA operon, intragenomic variation has the potential to mislead an investigator. For example, nonfunctional copies of ribosomal genes (i.e., pseudogenes) may be inadvertently assessed as "real" strain and/or species diversity if care is not taken. In this case, structural analysis has proven to be a powerful tool in identifying potential ITS2 pseudogenes, particularly when bacterial cloning is utilized, for exclusion from biodiversity estimates of dinoflagellates such as *Symbiodinium* (Thornhill et al., 2007). Such a system can now be extended to the PLDs with the models presented here. Presently, we are designing a web-based software utility that predicts the ITS2 secondary structure of dinoflagellates (including the PLDs) by a combination of the phylogenetic comparative method and minimum free energy (MFE; Mathews et al., 1999; Zuker, 2003). In the interim, we advise that novel PLD ITS2 sequences whose secondary structures differ significantly from our models should be interpreted cautiously and further screened as potential pseudogenes.

Acknowledgements

We thank Daniel Thornhill and Wayne Litaker for their helpful suggestions. This is contribution #40 to the Auburn University (AU) Marine Biology Program.[SS]

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.hal.2008.09.004.

References

Baldwin, B.G., 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the compositae. Mol. Phylogenet. Evol. 1, 3–16.

Burkholder, J.M., Glasgow Jr., H.B., 1997. *Pfiesteria piscicida* and other toxic *Pfiesteria*-like dinoflagellates: behavior, impacts, and environmental controls. Limnol. Oceanogr. 42, 1052–1075.

Burkholder, J.M., Gordon, A.S., Moeller, P.D., Law, J.M., Coyne, K.J., Lewitus, A.J., Ramsdell, J.S., Marshall, H.G., Deamer, N.J., Cary, S.C., Kempton, J.W., Morton, S.L., Rublee, P.A., 2005. Demonstration of toxicity to fish and to mammalian cells by *Pfiesteria* species: comparison of assay methods and strains. Proc. Natl. Acad. Sci. 102, 3471–3476.

Chen, C.A., Chang, C., Wei, N.V., Chen, C., Lein, Y., Lin, H., Dai, C., Wallace, C.C., 2004. Secondary structure and phylogenetic utility of the ribosomal internal transcribed spacer 2 (ITS2) in scleractinian corals. Zool. Stud. 43, 759–771.

Coleman, A.W., 2003. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. Trends Genet. 19, 370–375.

Coleman, A.W., 2007. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. Nucleic Acids Res. 35, 3322–3329.

Coleman, A.W., Mai, J.C., 1997. Ribosomal DNA ITS-1 and ITS-2 sequence comparisons as a tool for predicting genetic relatedness. J. Mol. Evol. 45, 168–177.

Coleman, A.W., Vacquier, V.D., 2002. Exploring the phylogenetic utility of ITS2 sequences for animals: a test case for abalone (*Haliotis*). J. Mol. Evol. 54, 246–257.

Connell, L.B., 2001. Rapid identification of marine algae (*Raphidophyceae*) using three-primer PCR amplification of nuclear internal transcribed (ITS) region from fresh and archived material. Phycologica 41, 15–21.

Furlong, J.C., Maden, B.E., 1983. Patterns of major divergence between the internal transcribed spacers of ribosomal DNA in *Xenopus borealis* and *Xenopus laevis*, and minimal divergence within ribosomal coding regions. EMBO J. 2, 443–448.

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, 1199–1206.

Gilbert, D.J., 1992. *loopDloop*, a Macintosh program for visualizing RNA secondary structure. Published electronically on the internet, available via anonymous ftp to ftp.bio.indiana.edu.

Goertzen, L.R., Cannone, J.J., Gutell, R.R., Jasen, R.K., 2003. ITS secondary structure derived from comparative analysis, implications for sequence alignment and phylogeny of the Asteraceae. Mol. Phylogenet. Evol. 29, 216–234.

Good, L., Intine, R.V.A., Nazar, R.N., 1997. Interdependence in the processing of ribosomal RNAs in *Schizosaccharomyces pombe*. J. Mol. Biol. 273, 782–788.

Gordon, A.S., Dyer, B.J., Seaborn, D., Marshall, H.G., 2003. Comparative toxicity of *Pfiesteria* spp., prolonging toxicity of *P. piscicida* in culture and evaluation of toxin(s) stability. Harmful Algae 1, 85–94.

Gordon, A.S., Dyer, B., 2005. Relative contribution of exotoxin and micropredation to icthyotoxicity of two strains of *Pfiesteria shumwayae* (Dinophyceae). Harmful Algae 4, 423–431.

Gottschling, M., Plötner, J., 2004. Secondary structure models of the nuclear internal transcribed spacer regions and 5.8S rRNA in Calciodinelloideae (Peridiniaceae) and other dinoflagellates. Nucleic Acids Res. 32, 307–315.

Gutell, R.R., Larsen, H., Woese, C.R., 1994. Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. Microbiol. Rev. 58, 10–26.

Hall, L.M.C., Maden, B.E.H., 1980. Nucleotide sequence through the 18S to 28S intergene region of a vertebrate ribosomal transcription unite. Nucleic Acids Res. 8, 5993–6005.

Hidenach, B.R., Stafford, D.W., 1984. Nucleotide sequence of the 18S-26S rRNA intergene region of the sea urchin. Nucleic Acids Res. 12, 1737–1747.

Hillis, D.M., Dixon, M.T., 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. Quart. Rev. Biol. 66, 411–453.

Hunter, R.L., LaJeunesse, T.C., Santos, S.R., 2007. Structure and evolution of the rDNA internal transcribed spacer (ITS) region 2 in the symbiotic dinoflagellates (*Symbiodinium*, Dinophyta). J. Phycol. 43, 120–128.

Jeong, H.J., Kim, J.S., Park, J.Y., Kim, J.H., Kim, S., Lee, I., Lee, S.H., Ha, J.H., Yih, W.H., 2005. *Stockeria algicida* n. gen., n. sp. (Dinophyceae) from the coastal waters off Southern Korea: morphology and small subunit ribosomal gene sequence. J. Eukaryot. Microbiol. 52 (4), 382–390.

Joseph, N., Krauskopf, E., Vera, M.I., Michot, B., 1999. Ribosomal internal transcribed spacer 2 (ITS2) exhibits a common core of secondary structure in vertebrates and yeast. Nucleic Acids Res. 27, 4533–4540.

Kimura, M., 1985. The role of compensatory neutral mutations in molecular evolution. J. Genet. 64, 7–19.

Kofoid, C.A., 1909. On *Peridinium steini* Jörgensen, with a note on the nomenclature of the skeleton of the Peridinidae. Arch. Protistenkd. 16, 25–47.

Litaker, R.W., Tester, P.A., Colorni, A., Levy, M.G., Noga, E.J., 1999. The phylogenetic relationship of *Pfiesteria piscicida*, Cryptoperidiniopsoid sp., *Amyloodinoum ocellatum*, and a *Pfiesteria*-like dinoflagellate to other dinoflagellates and apicomplexans. J. Phycol. 35, 1379–1389.

Litaker, W.R., Vandersea, M.W., Kibler, S., 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, 754–761.

Litaker, W.R., Steidinger, K.A., Mason, P.L., Landsberg, J.H., Shields, J.D., Reece, K.S., Haas, L.W., Vogelbein, W.K., Vandersea, M.W., Kibler, S.R., Tester, P.A., 2005. The

reclassification of *Pfiesteria shumwayae*: *Pseudopfiesteria*, gen. nov. J. Phycol. 41 (9), 643–651.
Litaker, W.R., Vandersea, M.W., Kibler, S., Reece, K.S., Stokes, N.A., Lutzoni, F.M., Yonish, B.A., West, M.A., Black, M.N.D., Tester, P.A., 2007. Recognizing dinoflagellate species using ITS rDNA sequences. J. Phycol. 43, 344–355.
Lu, L., Li, J., Cang, Y., 2002. PCR-based sensitive detection of medicinal fungi *Hericium* species from ribosomal internal transcribed spacer (ITS) sequences. Biol. Pharm. Bull. 8, 975–980.
Luo, G., Mitchell, T.G., 2002. Rapid identification of pathogenic fungi directly from cultures by using multiplex PCR. J. Clin. Microbiol. 40, 2860–2865.
Mai, J.C., Coleman, A.W., 1997. The internal transcribed spacer 2 exhibits a common secondary structure in green algae and flowering plants. J. Mol. Evol. 44, 258–271.
Marshall, H.G., Gordon, A.S., Seaborn, D.W., Dyer, B., Dunstan, W.M., Seaborn, A.M., 2000. Comparative culture and toxicity studies between the toxic dinoflagellate *Pfiesteria piscicida* and a morphologically similar cryptoperidiniopsoid dinoflagellate. J. Exp. Mar. Biol. Ecol. 255, 51–74.
Marshall, H.G., Hargraves, P.E., Burkholder, J.M., Parrow, M.W., Elbrachter, M., Allen, E.H., Knowlton, V.M., Rublee, P.A., Hynes, W.L., Egerton, T.A., Remington, D.L., Wyatt, K.B., Lewitus, A.J., Henrich, V.C., 2006. Taxonomy of *Pfiesteria* (Dinophyceae). Harmful Algae 5, 481–496.
Mason, P.L., Litaker, W.R., Jeong, H.J., Jeong, H.H., Reece, K.S., Stokes, N.A., Park, J.Y., Steidinger, K.A., Vandersea, M.W., Kibler, S., Tester, P.A., Vogelbein, W.K., 2007. Description of a new genus of *Pfiesteria*-like dinoflagellate, *Luciella* gen. nov. (Dinophyceae), including two new species: *Luciella masanensis* sp. nov. and *Luciella atlantis* sp. nov. J. Phycol. 43, 799–810.
Mathews, D.H., Sabina, J., Zuker, M., Turner, D.H., 1999. Expanded sequence dependence of thermodynamic parameters provides improved prediction of RNA secondary structure. J. Mol. Biol. 288, 911–940.
Mathews, D.H., Disney, M.D., Childs, J.L., Schroeder, S.J., Zuker, M., Turner, D.H., 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. Proc. Natl. Acad. Sci. 101, 7287–7292.
Michot, B., Joseph, N., Mazan, S., Bachellerie, J.P., 1999. Evolutionarily conserved structural features in the ITS2 of mammalian pre-rRNAs and potential interactions with the snoRNA U8 detected by comparative analysis of new mouse sequences. Nucleic Acids Res. 27, 2271–2282.
Moeller, P.D.R., Morton, S.L., Mitchell, B.A., Sivertsen, S.K., Fairey, E.R., Mikulski, T.M., Glasgow, H.B., Deamer-Melia, N.J., Burkholder, J.M., Ramsdell, J.S., 2001. Current progress in isolation and characterization of toxins isolated from *Pfiesteria piscicida*. Environ. Health Perspect. 109, 739–743.
Moeller, P.D.R., Beauchesne, K.R., Huncik, K.M., Davis, W.C., Christopher, S.J., Riggs-Gelasco, P., Gelasco, A.K., 2007. Metal complexes and free radical toxins produced by *Pfiesteria piscicida*. Environ. Sci. Technol. 41, 1166–1172.
Morgan, J.A., Blair, D., 1998. Trematode and monogenean rRNA ITS2 secondary structures support a four-domain model. J. Mol. Evol. 47, 406–419.
Noge, K., Mori, N., Tanaka, C., Nishida, R., Tsuda, M., Kuwahara, Y., 2005. Identification of astigmatid mites using the second internal transcribed spacer (ITS2) region and its application for phylogenetic study. Exp. Appl. Acarol. 35, 29–46.
Reddy, R., Rothblum, L.I., Subrahmanyam, C.S., Liu, M., Hennig, D., Cassidy, B., Busch, H., 1983. The nucleotide sequence of 8S RNA bound to preribosomal RNA of Novikoff hepatoma. J. Biol. Chem. 258, 584–589.
Savill, N.J., Hoyle, D.C., Higgs, P.G., 2001. RNA sequence evolution with secondary structure constraints: comparison of substitution rate models using maximum likelihood methods. Genetics 157, 399–411.
Schlötterer, C., Hauser, M.T., von Haeseler, A., Tautz, D., 1994. Comparative evolutionary analysis of rDNA ITS regions in *Drosophila*. Mol. Biol. Evol. 11, 513–522.
Schultz, J., Maisel, S., Gerlach, D., Muller, T., Wolf, M., 2005. A common core of secondary structure in the internal transcribed spacer 2 (ITS2) throughout the eukaryota. RNA 11, 361–364.
Seaborn, D.W., Tengs, T., Cerbin, S., Kokocinski, M., Marshall, H.G., 2006. A group of dinoflagellates similar to *Pfiesteria* as defined by morphology and genetic analysis. Harmful Algae 5, 1–8.
Steidinger, K.A., Burkholder, J.M., Glasgow, H.B.J., Hobbs, C.W., Garrett, J.K., Truby, E.W., Noga, E.J., Smith, S.A., 1996. *Pfiesteria piscicida* gen. et sp. nov. (Pfiesteriaceae fam. nov.), a new toxic dinoflagellate with a complex life cycle and behavior. J. Phycol. 32, 157–164.
Steidinger, K.A., Landsberg, J.H., Mason, P.L., Vogelbein, W.K., Tester, P.A., Litaker, R.W., 2006. *Cryptoperidiniopsis brodyi* gen. et sp. nov. (Dinophyceae), a small lightly armored dinoflagellate in the Pfiesteriaceae. J. Phycol. 42, 951–961.
Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (* and other methods). Version 4. Sinauer Associates, Sunderland, MA.
Tautz, D., Hancock, J.M., Webb, D.A., Tautz, C., Dover, G.A., 1988. Complete sequences of the rRNA genes of *Drosophila melanogaster*. Mol. Biol. Evol. 5, 366–376.
Taylor, F.J.R., 1999. Morphology (tabulation) and molecular evidence for dinoflagellate phylogeny reinforce each other. J. Phycol. 35, 1–6.
Thornhill, D.J., LaJeunesse, T.C., Santos, S.R., 2007. Measuring rDNA diversity in eukaryotic microbial systems: how intragenomic variation, pseudogenes, and PCR artifacts confound biodiversity estimates. Mol. Ecol. 16, 5326–5340.
Truby, E.W., 1997. Preparation of single-celled marine dinoflagellates for electron microscopy. Microsc. Res. Tech. 36, 337–341.
van der Sande, C.A.F.M., Kwa, M., van Nues, R.W., van Heerikhuizen, H., Raue, H.A., Plantar, R.J., 1992. Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisiae* ribosomal DNA. J. Mol. Biol. 223, 899–910.
van Nues, R.W., Rientjes, J.M.J., Morre, S.A., Mollee, E., Planta, R.J., Venema, J., Raue, H.A., 1995. Evolutionarily conserved structural elements are critical for processing of internal transcribed spacer 2 from *Saccharomyces cerevisiae* precursor ribosomal RNA. J. Mol. Biol. 250, 24–36.
Veldman, G.M., Klootwijk, J., van Heerikhuizen, H., Planta, R.J., 1981. The nucleotide sequence of the inter-genic region between the 5.8S and 26S rRNA genes of the yeast ribosomal RNA operon. Possible implications for the interaction between 5.8S and 26S rRNA and the processing of the primary transcript. Nucl. Acids Res. 9, 4847–4862.
Weekers, P.H.H., De Jonckheere, J.F., Dumont, H.J., 2001. Phylogenetic relationships inferred from ribosomal ITS sequences and biogeographic patterns in representatives of the genus *Calopteryx* (Insecta: Odonata) of the west Mediterranean and adjacent West European Zone. Mol. Phylogenet. Evol. 20, 89–99.
Wesson, D.M., 1992. Sequence and secondary structure comparisons of ITS rDNA in mosquitoes (Diptera: Culicidae). Mol. Phylogenet. Evol. 1, 253–269.
Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acid Res. 31, 3406–3415.