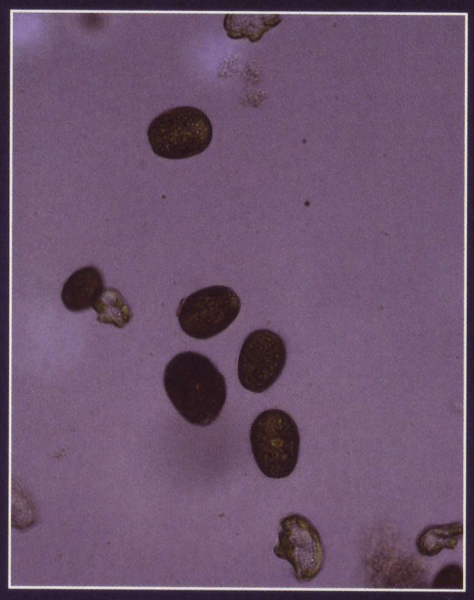
VOLUME 5 NUMBER 1

MARINE BIOTECHNOLOGY

An International Journal on the Molecular and Cellular Biology of Marine Life and Its Technological Applications

JANUARY/FEBRUARY 2003



Official Journal of:

- European Society for Marine Biotechnology
- Japanese Society of Marine Biotechnology
- Pan American Marine Biotechnology Association

1436-2228(200301/02)5:1;1-C

10126 ISSN 1436-2228

Springer

Mar. Biotechnol. 05, 1–12, 2003
DOI: 10.1007/s10126-002-0077-y

MARINE BIOTECHNOLOGY

Short Communication

Polyketide Synthase Genes from Marine Dinoflagellates

R.V. Snyder,[1] P.D.L. Gibbs,[2] A. Palacios,[1] L. Abiy,[2] R. Dickey,[3] J.V. Lopez,[4] and K.S. Rein[1,*]

[1]*Department of Chemistry, Florida International University, 11200 SW 8th St., Miami, FL 33199, U.S.A.*
[2]*University of Miami Rosenstiel School of Marine and Atmospheric Science, NIEHS Marine and Freshwater Biomedical Sciences Center, 4600 Rickenbacker Cswy, Miami, FL 33149, U.S.A.*
[3]*Gulf Coast Seafood Laboratory, U.S. Food and Drug Administration, Dauphin Island, AL 36528, U.S.A.*
[4]*Division of Biomedical Marine Research, Harbor Branch Oceanographic Institution, 5600 US 1 North, Ft. Pierce, FL 34946, U.S.A.*

Abstract: Rapidly developing techniques for manipulating the pathways of polyketide biosynthesis at the genomic level have created the demand for new pathways with novel biosynthetic capability. Polyketides derived from dinoflagellates are among the most complex and unique structures identified thus far, yet no studies of the biosynthesis of dinoflagellate-derived polyketides at the genomic level have been reported. Nine strains representing 7 different species of dinoflagellates were screened for the presence of type I and type II polyketide synthases (PKSs) by polymerase chain reaction (PCR) and reverse transcriptase PCR. Seven of the 9 strains yielded products that were homologous with known and putative type I PKSs. Unexpectedly, a PKS gene was amplified from cultures of the dinoflagellate *Gymnodinium catenatum*, a saxitoxin producer, which is not known to produce a polyketide. In each case the presence of a PKS gene was correlated with the presence of bacteria in the cultures as identified by amplification of the bacterial 16S ribosomal RNA gene. However, amplification from polyadenylated RNA, the lack of PKS expression in light-deprived cultures, residual phylogenetic signals, resistance to methylation-sensitive restriction enzymes, and the lack of hybridization to bacterial isolates support a dinoflagellate origin for most of these genes.

Key words: polyketide, polyketide synthase, toxin, dinoflagellate, biosynthesis.

INTRODUCTION

Polyketides are a highly diverse group of secondary metabolites, in terms of both structure and pharmacology, that share a common biogenic origin. The construction of the carbon framework is accomplished via the sequential Claisen condensations of acyl–coenzyme A (CoA) subunits. Thus the carbon chain assembly is related to fatty acid biosynthesis (O'Hagan, 1995); however, structural diversity in polyketides is achieved through the use of a larger pool of subunits and variations in the extent of post-condensation modifications (Cane, 1997).

The construction of all polyketides is directed by a family of enzymes, polyketide synthases (PKSs). Recently, polyketide biosynthetic genes from bacteria and fungi have been cloned, sequenced, and expressed in heterologous hosts. Examination of the genomic organization of PKS-encoding genes and genes for associated tailoring enzymes

Received June 20, 2002; accepted September 5, 2002.
**Corresponding author: telephone 305-348-6682; fax 305-348-3772; e-mail reink@fiu.edu*

has resulted in an unprecedented advancement in our understanding of the mechanisms of polyketide biosynthesis. Three classes of polyketide synthases have emerged (Gokhale and Dipika, 2001). Type I are large multifunctional enzymes in which several functional domains are located within a single protein. The functional domains of type I bacterial PKSs may be arranged into ''modules,'' with each module directing one round of chain extension (β-ketoacyl synthase [KS] domain) and postcondensation modifications. These are typically bacterial in origin and produce nonaromatic polyketides. Alternatively, fungal PKSs are architecturally similar to modular PKSs in that they are multifunctional enzymes; however, each is composed of a single module that is used iteratively. Fungal PKSs can produce either aromatic (Brown et al., 1996) or nonaromatic (Hendrickson et al., 1999) polyketides. Type II PKSs, multiprotein complexes of several individual enzymes, are found only in bacteria and usually code for aromatic polyketides. The individual enzymes of type II PKSs are used iteratively for each cycle of chain extension. Type III PKSs are found only in plants, utilize unusual starter units, and act on acyl-CoA thioesters. Alternatively, the substrates for type I and II PKSs are covalently linked to an acyl carrier domain or an acyl carrier protein (ACP), respectively.

The ability to functionally express biosynthetic genes has stimulated the development of a number of biosynthetic strategies for producing rare or novel polyketides in quantities. Combinatorial expression of cloned PKS enzymes or domains from type I and II PKSs have produced novel polyketides (McDaniel et al., 1993; Fu et al., 1994; Hutchinson and Fujii, 1995; Moore and Piel, 2000; Rodriguez and McDaniel, 2001). One disadvantage is the difficulty in predicting the structure of the final polyketide product a priori from the combinatorial expression of type II PKSs. In contrast, type I modular PKSs have been modified to produce novel polyketides in a highly predictable fashion, perhaps making the use of type I PKSs a more powerful approach to combinatorial design of novel polyketides. The emerging field of combinatorial biosynthesis has stimulated intense interest in identifying new sources of novel biosynthetic capability. These efforts are facilitated by the high degree of sequence homology within each class of PKS-encoding gene, particularly in the KS domain. The development of degenerate primer sets for highly conserved motifs of type I and type II PKSs has permitted the screening for these pathways in bacterial and fungal cultures, uncultivable bacterial strains or assemblages, and even soil (Seow et al., 1997; Bingle et al., 1999; Metsä-Ketelä et al., 1999).

Dinoflagellates are a highly diverse group of flagellated, unicellular protists. More than 4000 living and fossil species have been reported thus far. The majority are marine in origin and free-living, including benthic, epiphytic, and planktonic species, while others exist as endosymbionts with marine invertebrates, and some are parasitic (Taylor, 1987). Approximately 25 species of dinoflagellates are known to produce about 45 different polyketides (Rein and Borrone, 1999, and references cited therein). Dinoflagellate polyketides have often been associated with poisoning in humans who consumed tainted fish or shellfish, and this is typically how these compounds come to our attention (Plumley, 1997). Others have potential therapeutic value should the issue of limited supply be resolved (Bauer et al., 1995). Okadaic acid, a protein phosphatase inhibitor from *Prorocentrum lima*, has been used extensively as a biochemical probe to identify and characterize cellular processes that are regulated by phosphorylation and dephosphorylation (Fernandez et al., 2002). From the complex structures of dinoflagellate-derived polyketides, one would anticipate that the PKSs responsible for their biosynthesis would be similar to type I modular synthases. However, no PKS has been isolated from a dinoflagellate thus far, and no studies of the molecular basis of polyketide biosynthesis in dinoflagellates have been reported.

We have undertaken a screen of 7 species (9 strains) of cultured dinoflagellates for the presence of genes encoding type I PKS. Although we expected to amplify type I PKSs based on the structures of dinoflagellate-derived polyketides, molecular studies of polyketide biosynthesis have returned many surprises, and new paradigms emerge frequently. This and the lack of any information on PKSs from dinoflagellates prompted us to screen 4 dinoflagellate species for genes encoding type II PKS as well. We chose to screen by reverse-transcriptase polymerase chain reaction (RT-PCR) for a number of reasons. The size and potential complexity of some dinoflagellate genomes (Rizzo, 1987) could complicate the PCR reaction. Furthermore, RT-PCR ensures that the products are amplified from expressed genes, excluding pseudogenes and relic or unexpressed pathways. Finally, the possible presence of introns might complicate amplification or detection of the target sequence. In addition to indicating the expression of PKS capability, the amplified products will provide probes for library screens and for in situ hybridization to uncover the ultimate origins of the pathways.

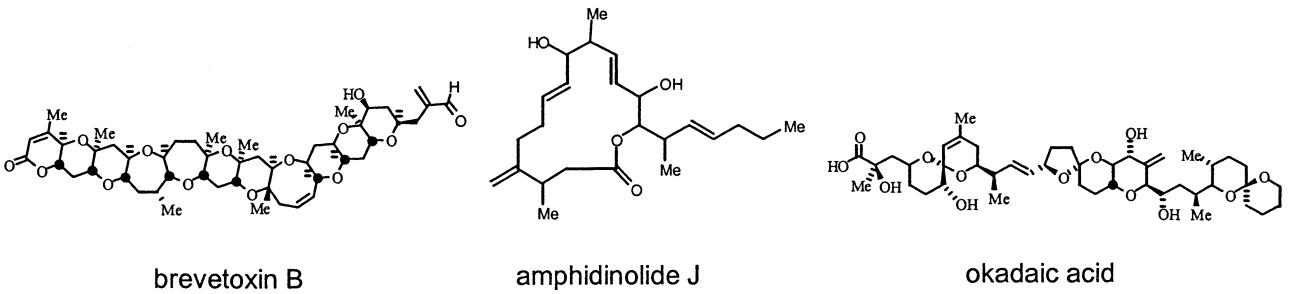


Figure 1. Representative structures of some of the polyketides produced by organisms in this study.

Three of the dinoflagellates examined are confirmed polyketide producers: the benthic dinoflagellates *Prorocentrum lima* and *Prorocentrum hoffmanianum* produce okadaic acid (Figure 1); the planktonic dinoflagellate *Karenia brevis* (formerly *Gymnodinium breve*), also known as the "Florida red tide" dinoflagellate, produces the polyether ladder brevetoxins. Four strains have not been established to produce polyketides, yet they belong to polyketide-producing genera. Three strains of *Amphidinium operculatum* (synonymous with *Amphidinium klebsii*) were examined. Strains of *A. klebsii* produce macrolides (amphidinolides) and polyhydroxypolyenes (Kobayashi and Ishibashi, 1993). One strain of *Symbiodinium* was examined. Again, this particular strain is not known to produce a polyketide. However, others are known producers of the zooxanthellatoxins (Nakamura et al., 1995). Finally, 2 species not known for polyketide production, *Gymnodinium catenatum* and *Amphidinium carterae*, were examined.

Materials and Methods

Cultures and Culturing Conditions

P. lima, *K. brevis*, and *G. catenatum* were obtained from the culture collection of the University of Miami's Marine and Freshwater Biomedical Sciences Center. *P. hoffmanianum* was obtained from the Gulf Coast Seafood Laboratory, U.S. Food and Drug Administration, Dauphin Island, Ala. *A. operculatum* (strains CCMP120, 121, and 1342), *A. carterae* (CCMP1314), and *Symbiodinium* sp. (CCMP831) were obtained from the Provasoli-Guillard Center for the Culture of Marine Phytoplankton, CCMP (Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, Me.). Cultures were maintained in Wilson's NH-15 media (*K. brevis*), f/2 media (*P. lima*, *P. hoffmanianum*, *A. operculatum*, *A. carterae*, and *G. catenatum*), and Prov.50 (*Symbiodinium*) at 20° to 22°C under constant illumination from either Cool White or Grow-Lux wide-spectrum lamps (Guillard, 1975).

Isolation of RNA and DNA

Dinoflagellate cells were isolated by centrifugation (3000 *g*, 5 minutes). DNA was purified by a modified proteinase K digestion (Sambrook et al., 1989). Pellets (typically 100 μl) were suspended in a 5-fold volume of digestion solution (50 mM Tris, 50 mM EDTA, pH 8.0, 200 mM NaCl, 1% sodium dodecylsulfate [SDS], 0.1% proteinase K), incubated overnight at 55°C, followed by a phenol–chloroform–isoamyl alcohol (25:24:1) extraction. DNA was precipitated with isopropanol, washed with 70% ethanol, and resuspended in TE (pH 8). Total RNA was purified by the method described for *Gonyaulax polyhedra* (Bae and Hastings, 1994), quantified spectrophotometrically, and shown by gel electrophoresis to be of high quality by virtue of high molecular weight and sharp ribosomal RNA bands. The total RNA was treated with DNAse to remove any residual DNA.

RT-PCR and Genomic PCR

The isolated RNA was reverse-transcribed by M-MuLV reverse transcriptase (Promega) using random, degenerate, or sequence-specific primers. PCR was performed using the generated complementary DNA or DNA as a template and 4 different primer sets (Life Technologies) (Table 1): the 2 pairs of degenerate PKS primers (PKSI and PKSII); bacterial 16S ribosomal primers; and universal eukaryotic 18S ribosomal primers (as a positive control). *Taq* DNA polymerase was purchased from either Promega or Eppendorf. A variety of protocols were used to generate products ("standard," touchdown, low-annealing); however, typical conditions were as follows: "hot-to-80 start" (95°C/5 min, 80°C/add *Taq* DNA polymerase), followed by a touch

Table 1. Primers Sets Used to Amplify PKS or rRNA Genes

Primer	Sequence 5′-3′	Origin
PKSI forward	MGIGARGCIYTICARATGGAYCCICARMG	Brown et al., 1996
PKSII reverse	GGRTCNCCIARYTGIGTICCIGTICCRTGIGC	
PKSII forward	TSGCSTGCTTCGAYGCSATC	Metsä-Ketelä et al., 1999
PKSII reverse	TGGAANCCGAABCCGCT	
18S rRNA forward	GGTTGATCCTGCCAGTAGTCATATGCCTG	Rowan and Powers 1992
18S rRNA reverse	GATCCTTCCGCAGGTTCACCTACGGAAACC	
16S rRNA forward	GGAGAGTTTGATCMTGGCT	Prokic et al., 1998
16S rRNA reverse	CCAGGGTATCTAATCCTGTT	

down protocol, which lowered the annealing temperature from 65°C to 55°C over 10 cycles, denaturation (95°C/30 s), annealing (65°C to 55°C, decreasing by 1°C per cycle/ 1.5–2 min), extension (72°C/2–3 min), for 10 cycles. This was followed by 29 cycles of denaturation (95°C/30 s), annealing (55°C/1.5–2 min), extension (72°C/2–3 min). A final extension at 72°C for 10 minutes completed the reaction. PCR was performed on a Mastercycler Gradient thermal cycler (Eppendorf).

Cloning and Sequencing of Amplified Fragments

The anticipated DNA fragments from the RT-PCR reactions using PKS (700 bp for PKSI) primers were purified from a 0.8% low-melt agarose gel and ligated into the T/A site of the pCR2.1 plasmid vector (Invitrogen) with T4 DNA ligase. Competent *Escherichia coli* (INVαF′) were transformed according to the manufacturer's instructions. For *K. brevis* 80 different clones were tested for insert size, restriction site polymorphism, and hybridization with the original PCR band. By this analysis, 15 unique clones were chosen for sequencing. For the other species, roughly 40 clones were tested for insert size, and 3 to 8 clones of the appropriate size were chosen at random for sequencing. Both strands were sequenced on an ABI automated sequencer (PerkinElmer) using the M13R and T7 primers.

Southern Analysis

Southern blots were prepared on Duralon-UV membranes (Stratagene) via capillary rise (Sambrook et al., 1989) or Posiblot pressure blotting (Stratagene) according to the manufacturer's instructions. Probe DNA was labeled with [α-^{32}P]dCTP (New England Nuclear) using random primers and the Klenow fragment (New England Biolabs) and purified with NucTrap (Stratagene) columns. Southern blots were screened according to established procedures (Sambrook et al., 1989).

Sequence Analysis

The translated protein sequence was compared with protein databases over the Internet using the program Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990, 1999). All sequences were aligned with CLUSTAL W (Thompson et al., 1994) and then checked by eye. The alignments used for phylogenetic reconstructions are available from the authors upon request. Phylogenetic analyses of KS and KS-like sequences were performed using PAUP 4.0b10 (PPC) (Swofford, 2000), and included (a) RANDOM TREES test to determine phylogenetic signal (b) base composition and codon biases. Owing to the high sequence divergence among KS sequences both among and within species, distance methods applied to amino acid sequences were mainly relied upon to construct phylogenetic dendograms. For the same reason, only ingroup (dinoflagellate) DNA sequences were compared. Bootstrap resampling (with a minimum of 500 replicates) was performed to assess statistical support for any reconstructed clades (Hillis et al., 1996).

To compare dinoflagellate KS sequences with known KS metabolite biosynthetic seqeunces, previously characterized KS (or fatty acid synthase [FAS] sequences were retrieved from GenBank and included in the analysis: Apicomplexan (*Cryptosporidium parvum*, AF082993), *Microcystis aeruginosa* (AAF00959), *Bacillus subtilis* (NP_389601), *Stigmatella aurantiaca* (AAK57187), *Aspergillus terreus* (AF141925), *Cochliobolus heterostrophus* (AAB08104), DEBS-6 for erythromycin (M63677), rifamycin (AF040570), and rapamycin (X86780). Tree rooting was initially performed using midpoint rooting, which places the root at the longest branch of the tree. Alterna-

Table 2. RT-PCR Results[a]

Organism	PKSI	PKSII	16S
P. lima	+	+	+
P. hoffmanianum	+	−	+
K. brevis	+	−	+
Symbiodinium sp.	+	−	+
A. operculatum			
CCMP1342	+	ND	+
CCMP120	+	ND	+
CCMP121	−	ND	−
A. carterae	−	ND	−
G. catenatum	+	ND	+

[a]ND, not determined. [b]Several small (≤200 bp) bands were generated, but none of the expected size (600 bp).

tively, outgroup sequences such as type II KS genes (granaticin, P16540;and actinorhodin, Q02062), and a FAS (Fix-23, X64131; chicken FAS, P12276) were also used in some reconstructions. Other putative KS-like sequences from marine bacteria were included for comparison (J.V. Lopez, unpublished work).

Isolation of Bacteria from Cultures

Bacteria were subcultured from *P. lima* and *K. brevis* by inoculating agar plates with aliquots of dinoflagellate culture. A variety of media were used including YG plus f/2 vitamin mixture, YG alone, and Difco Marine Agar. The plates were incubated at 25°C for 3 to 7 days. Single colonies displaying different morphologies were replated several times to ensure homogeneity of the isolates. Bacterial assemblages were collected from lawns. Genomic DNA was purified from isolated strains or bacterial assemblages by proteinase K digestion/CTAB extraction (Sambrook et al., 1989).

Results and Discussion

Two primer sets were selected to screen for PKS-encoding genes (Table 1). The criteria for selection of the primer sets were as follows: First, the gene product should be a required element of PKSs. Second, since we were screening by RT-PCR, the 2 priming sites should be located on a single transcript; for instance, primers spanning multiple genes encoding type II PKSs (Seow et al., 1997) would not be appropriate. Finally, the primer sets should flank at least one highly conserved motif to help identify the product after sequencing.

Type II PKSs must contain a set of 3 genes known as the ''minimal PKS.'' These include 2 KS subunits, KS_α and KS_β (sometimes referred to as the ''chain length factor'' or CLF), and the acyl carrier protien (ACP) to which the growing polyketide chain is covalently attached. The PKSII primers were designed to amplify a portion of the highly conserved KS_α subunit (Metsä-Ketelä et al., 1999). The type I PKS primer set (PKSI) was also based on the highly conserved KS domains of 17 bacterial and fungal type I PKSs as well as the rat FAS (Brown et al., 1996).

Using the KS_α primer set (PKSII), a band of 600 bp was anticipated and was produced with DNA containing the *act* (actinorhodin) PKS (Fernandez-Moreno et al., 1992) as a template. Among the 4 dinoflagellates screened, *K. brevis*, *P. lima*, *P. hoffmanianum*, and *Symbiodinium* only *P. lima* cDNA yielded any products (Table 2). Several small bands of less than 200 bp were observed. None of these bands were of the anticipated size, and they were not sequenced.

Seven of the 9 dinoflagellates, representing six of seven species examined, yielded RT-PCR products of the anticipated size when using the type I PKS primer set (Table 2). PCR reactions using *A. operculatum* (CCMP121) and *A. carterae* (CCMP1314) cDNA as a template with the degenerate PKS primers did not yield the expected 700-bp band. RT-PCR was compared to PCR using genomic DNA as a template (representative examples in Figure 2). While *P. lima* and *G. catenatum* did provide a product of the expected size from both the RT and the genomic PCR, *K. brevis*, *A. operculatum* (CCMP1342), *P. hoffmanianum*, and *Symbiodinium* sp. yielded a band of the anticipated size

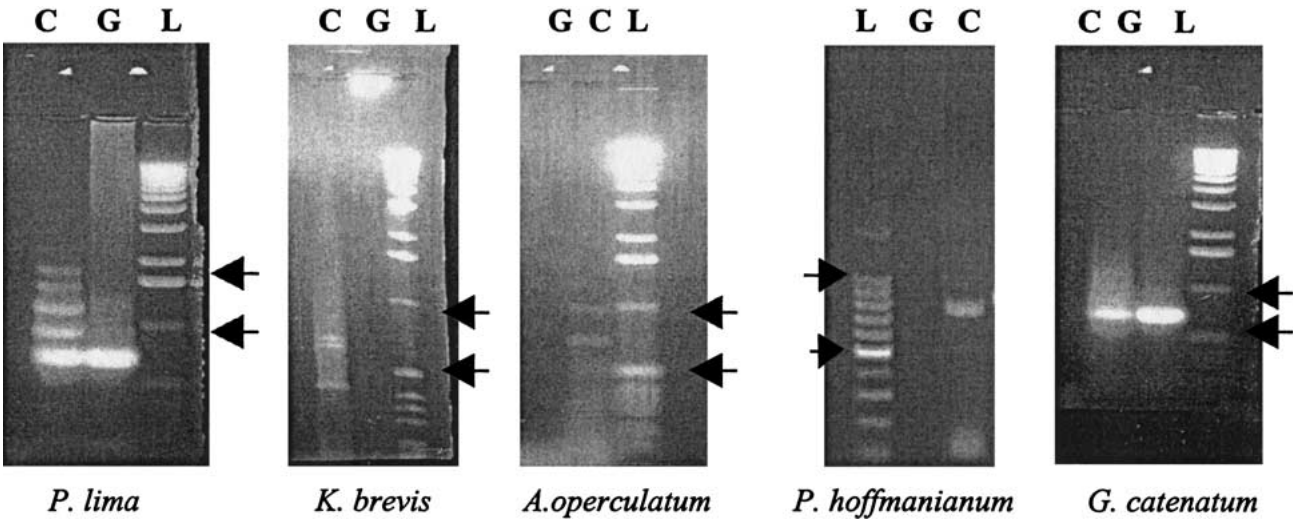


Figure 2. Comparison of genomic PCR (G lanes) and RT-PCR (C lanes) for 5 dinoflagellates using degenerate type I PKS primers. L lanes, DNA ladder; 500-bp and 1-kbp bands are indicated with arrows.

only in the RT-PCR. This could have been due to the presence on introns, the complexity of the genomic templates, or the occurrence of contaminants in the genomic DNA that interfered with the polymerase. Cloning and sequencing of the gene fragments from the RT-PCR reactions revealed products with high homologies to known and putative type I PKSs (Table 3) when submitted for BLAST analysis (Altshul et al., 1990). Of the dinoflagellates that yielded a 700-bp band, only *K. brevis* yielded *some* amplified products that were not PKS related. Eleven of 15 sequenced gene fragments from *K. brevis* were either related to photosystem I or II genes or provided no match in the BLAST search. The remaining 4 were PKS related.

Although they are probably the most conserved among all of the different catalytic domains that comprise a typical PKS genetic locus, pairwise amino acid distances between KS sequences still average about 55%, and reach beyond 65% divergence. By contrast, intraspecific (or intralocus) KS divergences appear to be consistently less (J.V. Lopez, personal communication) in the range of 10%–35% divergence, which is also reflected in the high bootstrap values for module (mod) clusters within the same locus for terrestrial actinomycetes. In this context we would expect dinoflagellate KS sequences to follow similar patterns. In the amino acid distance matrix and bootstrap neighbor-joining tree shown in Figure 3, there is relatively high identity (86%) between *K. brevis* AT2-10L and *K. brevis* AT2-15 PKS sequences, unlike that seen among most pairwise KS sequence comparisons, which suggests that these 2 sequences are derived from the same or a recently diverged locus. However, the alignments also showed that the active site cysteine (of the CSSsL consensus) was mutated in *K. brevis* AT2-10L, indicating that this may be a recently created pseudogene.

The neighbor-joining tree also indicated that none of the newly characterized dinoflagellate sequences exhibited either monophyly or strong clustering with previously characterized and empirically determined fungal or actinomycete KS sequences. For example, *P. lima* sequences were scattered at various locations of the tree, or suggesting high substitution rates, or long divergence times, or both. Many *G. catenatum* sequences behaved in the same manner, failing to associate into a well-supported single clade. For example, *G. catenatum* 23L grouped strongly with the *Symbiodinium* W9 (see more below) and with a *Bacillus subtilis* KS, both of which lack confirmed polyketide metabolite production. In contrast, affinity between most *K. brevis* (except *K. brevis* AS 1-1L) and *Cryptosporidium parvum* KS sequences was evident in the 63% bootstrap support for this clade. *Cryptosporidium parvum* belongs to Coccidia, a sister clade to the Dinophyceae within the Alveolata. One *P. lima* sequence (*P lima* 28L) tended to fall away from the large KS clade (99% bootstrap support) and cluster more with FAS sequences, exhibiting only 44% divergence from Fix-23.

Interestingly, 2 pairs of KS sequences unexpectedly appeared to be nearly identical; both *P. lima* 5L with *A. operculatum* 120 WB, and *Symbiodinium* W9 with G. *catenatum* 23L, shared greater than 99.0% amino acid identities with each other, respectively. This finding is

Table 3. BLAST Analysis of KS clones

Dinoflagellate	Clone	Highest Homology	E-value	Accession
A. operculatum 120	WB	putative type I polyketide synthase WcbR [*Burkholderia mallei*]	2.00E-84	AF285636
A. operculatum 120	WJ	modular polyketide synthase [*Streptomyces avermitilis*]	3.00E-36	BAB69192
A. operculatum 1342	AK112W	MtaB [*Stigmatella aurantiaca*]	1.00E-67	AF188287
G. *catenatum*	9L	putative type I polyketide synthase WcbR [*Burkholderia mallei*]	2.00E-61	AF285636
G. catenatum	19L	epoD [P*olyangium cellulosum*]	1.00E-70	AF217189
G. catenatum	23L	polyketide synthase [*Bacillus subtilis*]	5.00E-55	CAB13603
K. brevis	AS1-1L	polyketide synthase [*Nostoc* sp. PCC7120]	8.00E-52	BAB74379
K. brevis	AT2-15	type I fatty acid synthase homolog - *Cryptosporidium parvum*	8.00E-44	AAC99407
K. brevis	AT1-6L	polyketide synthase [*Nostoc* sp. PCC7120]	2.00E-46	BAB74379
K. brevis	AT2-10L	type I fatty acid synthase homolog - *Cryptosporidium parvum*	6.00E-41	AAC99407
P. lima	2L/5L/14L	putative type I polyketide synthase WcbR [*Burkholderia mallei*]	8.00E-84	AF285636
P. lima	9W/16L/25L	polyketide synthase pksE [*Clostridium acetobutylicum*]	3.00E-64	AAK81287
P. lima	28L	FATTY ACID SYNTHASE TRANSMEMBRANE PROTEIN [*Sinorhizobium meliloti*]	5.00E-62	CAC45149
P. lima	4L	polyketide synthase [*Nostoc* sp. PCC7120]	3.00E-76	BAB78014
Symbiodinium spp.	W9	polyketide synthase [*Bacillus subtilis*]	5.00E-54	CAB13603

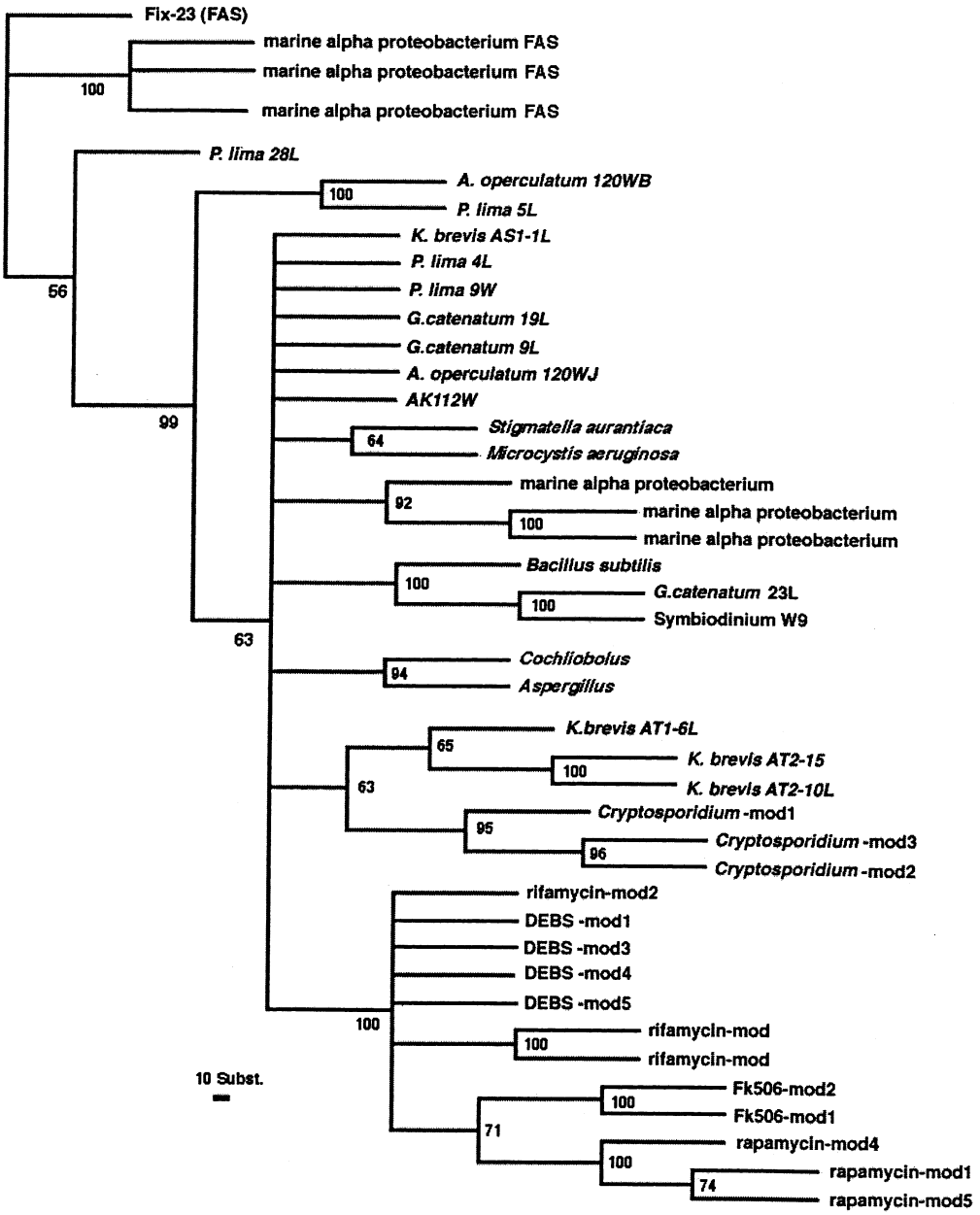


Figure 3. Phylogenetic tree based on KS amino acid sequences generated with the neighbor-joining algorithm in PAUP. Bootstrap percentages for reconstructed clades are displayed at each node. KSs from several distinct modules (mod) are included from well-characterized terrestrial actinomycetes. The marine α-proteobacteria KS-like sequences have not been empirically confirmed, but displayed high identities (>50%) to previously determined KS sequences in GenBank. Although not shown in this tree, inclusion of type II sequences (from granaticin and actinorhodin loci) always clustered as a separate clade from type I sequences, further supporting their distinctiveness.

supported at the DNA level, which similarly showed only 2.1% and 1.7% Kimura 2N corrected distances between *P. lima* 5L and *A. operculatum* 120WB and between *Symbiodinium* W9 and *G. catenatum* 23L, respectively. These results suggest that within the 2 pairs, the sequences are virtually identical. Cross hybridization of KS probes to genomic DNA, under high-stringency conditions, within these 2 pairs of KS sequences confirms this observation.

Overall, the large distances among most KS sequences involved in the analysis made reliable reconstruction of the actual branching order difficult, resulting in the large polytomy of the tree in Figure 3. The dinoflagellate KS sequences do not have a clear phylogenetic origin, being interspersed with bacterial and fungal type I KS sequences; alternatively, dinoflagellate KS sequences may be so ancient that the relationships have been obscured.

The 16S ribosomal primers were used to assess for the presence of bacteria (Prokic et al., 1998). Since toxic dinoflagellate cultures have not been successfully maintained axenically, the ultimate origins of the toxins are uncertain. The existence of toxic and nontoxic strains of dinoflagellates and the sudden unexplained loss of toxicity of dinoflagellates in culture has led to speculation that associated bacteria are responsible for toxin biosynthesis (Tosteson et al., 1989; Taylor, 1993). However, no polyketide-producing bacteria have been isolated from a dinoflagellate

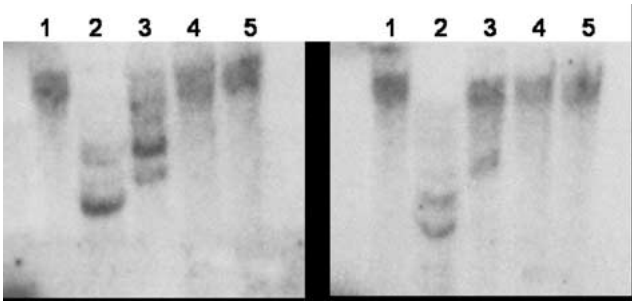


Figure 4. *K. brevis* restriction digests probed with *K. brevis* probes AS1-1L1L (left) and AT2-10L (right). Lane 1 *Hpa*II; lane 2, *Msp*I; lane 3, *Hae*III; lane 4, *Eco*R1; lane 5, uncut.

culture in spite of significant efforts by several groups to do so. The bacteria associated with *P. lima* have received considerable scrutiny, and conflicting reports in the literature on the origins of okadaic acid can be found (Rausch de Traubenberg and Lassus, 1991). Two groups have since used immunolocalization techniques to localize OA to the dinoflagellates (Zhou and Fritz, 1994; Rausch de Traubenberg et al., 1995). However, in our hands, the presence of PKS-encoding genes correlated with the presence of bacteria. *A. carterae* and *A. operculatum* (CCMP 121) failed to yield an RT-PCR product with either the PKSI primer set or the 16S ribosomal primers. All other dinoflagellates examined produced a band using either the PKS primers or the 16S rRNA primers. While correlation does not imply causation, the results of the PCR do not rule out a bacterial origin for these biosynthetic genes. The phylogenetic tree also does not obviate this possibility.

However, Southern analysis of *K. brevis* and *P. lima* DNA as well as DNA from a total of 24 bacterial isolates supports a dinoflagellate origin for the PKS-encoding genes. Intense hybridization of PKS probes was observed only to the genomic dinoflagellate DNA, indicating that the PKS genes originated from an organism making up a substantial proportion of the overall population and not from a minor contaminant. In no instance was hybridization observed to genomic DNA from a bacterial isolate or bacterial assemblage.

Dinoflagellate DNA is refractory to many restriction enzymes (Lee et al., 1993). This is no doubt due to the high degree of methylation and the substitution of hydroxymethyluracil for thymine. Restriction digests of *K. brevis* DNA were probed with PKS gene fragments as shown in Figure 4. The most notable result is the difference in effectiveness of the pair HpaII and MspI in the restriction of *K. brevis* DNA. These 2 enzymes are isoschizomers (both recognize the sequence CCGG). However, *Hpa*II is blocked by methylation at either cytosine, whereas *Msp*I is blocked only by methylation of the external cytosine. The observation that *Msp*I is able to restrict *K. brevis* DNA, but *Hpa*II is not, is indicative of heavy methylation of the target (*PKS*) sequence, at the dinucleotide CG, which is typical of eukaryotic methylation. Furthermore, *Hpa*II was ineffective in cutting *P. lima* DNA as well. The enzyme *Hae*II (GGCC) is sensitive only to methylation of the external cytosine and was also found to be effective in cutting *K. brevis* DNA. Among the restriction enzymes tested with *P. lima* DNA, only *Apa*I was effective. *Apa*I recognizes the sequence GGGCCC, and is insensitive to internal methylation. Other enzymes, *Eco*RI (GAATTC), *Bam*HI (GGATTC), *Xho*I (CTCGAG), *Kpn*I (GGTACC,) *Pst*I (CTGCAG), and *Hind*III (AAGCTT), all contain adenine or thymine in their recognition sequence, and thus would be expected to be inhibited by the substitution of thymine by hydroxymethyluracil. None were effective in cutting the target sequence.

Cultures of *P. lima* deprived of light for 3 weeks were screened for the expression of 16S, 18S, PKS gene expression. Microscopic observation of the cultures indicated that no viable dinoflagellate cells were present. Only the 16S ribosomal gene was amplified by RT-PCR from these light-deprived cultures (Figure 5) but not the 18S or PKS genes. Again, these findings strongly implicate a dinoflagellate origin for PKS-encoding genes. Finally, the use of oligo $(dT)_{15}$ the primer for first-strand cDNA synthesis followed by PCR with degenerate PKS primers should yield PKS products if the genes are polyadenylated (i.e., eukaryotic). Using *P. lima* RNA as the template, this sequence yielded an expected product of 700-bp (Figure 6). The 700 bp PCR band was shown by Southern analysis to be PKS related.

The present study indicates that molecular methods can be used to detect PKS genes from marine dinoflagellates and that type I PKS genes may be widespread among dinoflagellates. Our results suggest that polyketides derived from the organisms tested are produced by type I PKSs. If type II PKSs exist in dinoflagellates, they are refractory to the primers used or are in untested species. The presence of PKS-encoding genes in *G. catenatum* was unexpected, but demonstrates the potential to identify novel polyketides from cultured dinoflagellates. While we cannot rule out a bacterial origin for these genes, we found no evidence to support it aside from their co-occurrence with bacteria. Codon usage and phylogenetic position for the amplified gene fragments were not consistent with *Streptomycetes*, the most common producers of bacterial type I polyketides.

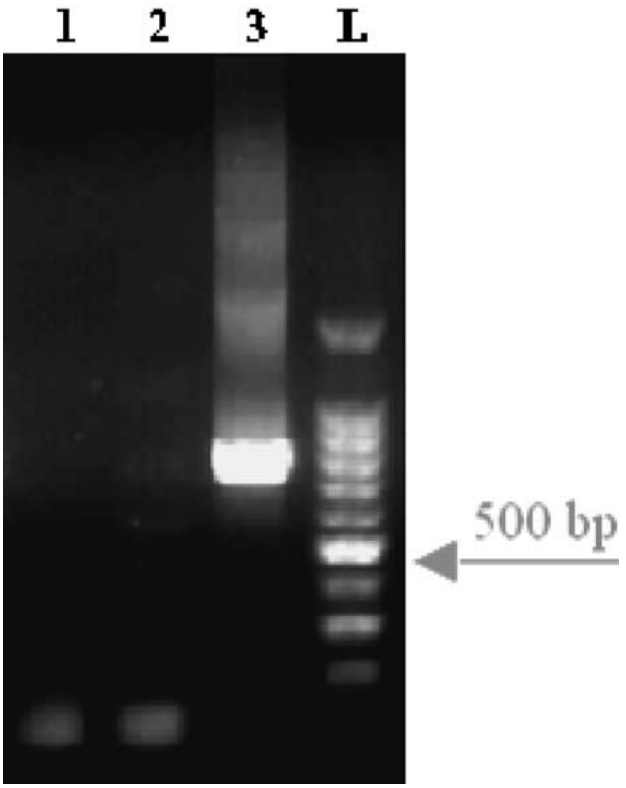


Figure 5. RT-PCR using cultures of *P. lima* deprived of light for 3 weeks. Lane 1, PKS primers; lane 2, 185 primers; lane 3, 165 primers; lane L, loc bp ladder.

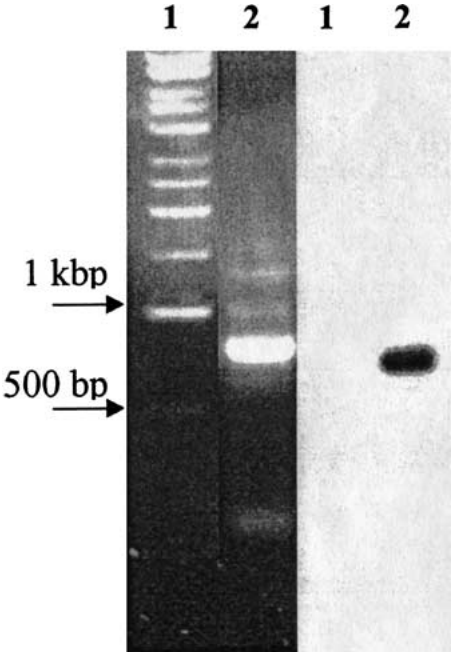


Figure 6. RT-PCR using oligo$(dT)_{15}$ the primer for the first-strand cDNA synthesis. **Left:** Lane 1, DNA ladder; lane 2, PCR reaction using *P. lima* cDNA as the template. **Right:** Southern blot of gel on left probed with heterologous *P. lima* PKS probes.

The homology between several dinoflagellate KS sequences and those from *Crytosporidium parvum*, the resistance of these genes to many restriction enzymes, the lack of hybridization of the amplified products to bacterial assemblages or isolates, and the absence of PKS gene expression in light-deprived cultures strongly support a dinoflagellate origin for at least *some* of these KS genes.

Marine microorganisms have long been a rich source of biologically active secondary metabolites (Faulkner, 2002). The ability to manipulate secondary metabolic pathways has increased interest in marine microorganisms as sources with novel biosynthetic capabilities. The uniqueness of dinoflagellates-derived polyketides, the unusual labeling patterns exhibited in feeding experiments, and the extensive tailoring reactions make their biosynthetic pathways extremely attractive targets, as coupling of these pathways with those of bacterial origin offers unprecedented biosynthetic potential. Furthermore, each new pathway provides insights into the construction of these highly complex molecules and the cooperativity between enzymes, and a fundamental understanding of toxin biosynthesis could lead to mechanisms for detection, mitigation, or control of these environmentally significant organisms and their toxins.

Acknowledgments

This work was supported by NIH R21 GM60012-01, USDA NRICGP 95-37201-1993, NIH P30 ES05705. R.V.S. was supported by GM 61347. J.V.L. was supported by NSF-MCB 9974984.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol* 215:403–410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* 25:3389–3402.

Bae, Y.M., and Hastings, J.W. (1994). Cloning, sequencing and expression of dinoflagellate luciferase DNA from a marine algae, *Gonyaulax polyhedra*. *Biochim Biophys Acta* 1219:449–456.

Bauer, I., Maranda, L., Young, K.A., Shimizu, Y., Fairchild, C., Cornell, L., MacBeth, J., and Huang, S. (1995). Isolation and

structure of caribenolide I, a highly potent antitumor macrolide from a cultured free-swimming dinoflagellate, *Amphidinium sp.* S1-36-5. *J Org Chem* 60:1084–1085.

Bingle, L.E., Simpson, T.J., and Lazarus, C.M. (1999). Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. *Fungal Genet Biol* 26:209–223.

Brown, D.W., Yu, J.H., Kelkar, H.S., Fernandes, M., Nesbitt, T.C., Keller, N.P., Adams, T.H., and Leonard, T.J. (1996). Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. *Proc Natl Acad Sci USA* 93:1418–1422.

Cane, D. (1997) Introduction: Polyketide and Non-ribosomal Polypeptide Biosynthesis. From Collie to Coli Biosynthesis. *Chem Rev* 97:2463–2464.

Carrasquero-Verde, J.R. (1999). Role of associated bacteria in *Heterosigma carterae* toxicity to salmonids. *Aquat Toxicol* 45:19–34

Faulkner, D.J. (2001). *Nat Prod Rep* 19:1–49.

Fernandez, J.J., Candenas, M.L., Souto, M.L., Trujillo, M.M., and Norte, M. (2002). Okadaic acid. Useful tool for Studying Cellular processes. *Cur Med Chem* 9:229–262.

Fernandez-Moreno, M.A., Martinez, E., Boto, L., Hopwood, D.A., and Malpartida, F. (1992). Nucleotide sequence and deduced functions of a set of co-transcribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. *J Biol Chem* 267:19278–19290.

Fu, H., McDaniel, R., Hopwood, D.A., and Khosla, C. (1994). Engineered biosynthesis of novel polyketides: stereochemical course of two reactions catalyzed by a polyketide synthase. *Biochemistry* 33:9321–9326.

Gokhale, R.S., and Dipika, T. (2001). Biochemistry of polyketide synthases In: *Biotechnology* 2nd ed, Hans-Juergen, R., ed. Weinheim, Germany: Wiley-VCH Verlag GmbH, 10:341–372.

Guillard, R.R.L. (1975). Culture of phytoplankton for feeding marine invertebrates. In: *Culture of Marine Invertebrate animals,* Smith, W.L., and Chanley, M.H. eds. New York N.Y.: Plenum Press, 26–60.

Hallegraeff, G.M. (1993). *Phycologia* 32:79–99.

Hendrickson, L., Davis, C.R., Roach, C., Nguyen, D.K., Aldrich, T., McAda, P.C., and Reeves, C.D. (1999). Lovastatin biosynthesis in *Aspergillus terreus:* characterization of blocked mutants, enzyme activities and a multifunctional polyketide synthase gene. *Chem Biol* 6:429–439.

Hillis, D.M., Moritz, C., and Mable, B.K., eds. (1996). *Molecular Systematics*. 2 ed. Sunderland, mass.: Sinaeur.

Hutchinson, C. R., and Fujii, I. (1995). Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. *Ann Rev Microbiol* 49:201–238.

Kobayashi, J., and Ishibashi, M. (1993). Bioactive metabolites of symbiotic marine microorganisms. *Chem Rev* 93:1753–1769.

Lee, H.L., Mittag, M., Sczekan, S., Morse, D., and Hastings, W. (1993). Molecular cloning and genomic organization of a gene for luciferin-binding protein from the dinoflagellate *Gonyaulax polyhedra*. *J Biol Chem* 268:8842–8850.

McDaniel, R., Ebert-Khosla, S., Hopwood, D.A., and Khosla, C. (1993). Engineered biosynthesis of novel polyketides. *Science* 262:1546–1550.

Metsä-Ketelä, M., Salo, V., Halo, L., Hautala, A., Hakala, P.M., and Ylihonko, K. (1999). An efficient approach for screening minimal PKS genes from *Streptomyces*. *FEMS Microbiol Lett* 180:1–6.

Moore, B.S., and Piel, J. (2000). Engineering biodiversity with type II polyketide synthase genes. *Antonie Van Leeuweenhoek* 78:391–398.

Nakamura, H., Asari, T., Fujimaki, K., Maruyama, K., Murai, A. (1995). Zooxanthellatoxin-B, vasoconstrictive congener of zooxanthellatoxin-A from a symbiotic dinoflagellate *Symbiodinium* sp. *Tet Lett* 36:7255–7258.

O'Hagan, D. (1995). The Biosynthesis of Fatty Acid and Polyketide Synthesis. *Nat Prod Rep* 12:1–32.

Plumley, F.G. (1997). Marine algal toxins: biochemistry, genetics and molecular biology. *Limnol Oceanogr* 42:1252–1264.

Prokic, I, Brummer, F., Brigge, T., Gortz, H.H., Gerdts, G., Schutt, C., Elbrachter, M., and Muller, E.G. (1998). Bacteria of the genus *Roseobacter* associated with the toxic dinoflagellate *Prorocentrum lima*. *Protist* 149:347–357.

Rausch de Traubenberg, C., and Lassus, P. (1991). Dinoflagellate toxicity: Are marine bacteria involved Evidence from the literature. *Mar Microb Food Webs* 5:205–226.

Rausch de Traubenberg, C., Soyer-Gobillard, M.-O., Geraud, M.-L., and Albert, M. (1995). The toxic dinoflagellate *Prorocentrum lima* and its associated bacteria, II:imunolocalization of okadaic acid in axenic and non-axenic cultures. *Eur J Protistol* 31:383–388

Rein, K.S., and Borrone, J. (1999). Polyketides from dinoflagellates: origins, pharmacology and biosynthesis. *Comp Biochem Physiol B* 124:117–131.

Rizzo, P.J. (1987). Biochemistry of the dinoflagellate nucleus. In: *The Biology of Dinoflagellates*, Taylor, F.J.R., Ed. Malden, Mass.: Blackwell Scientific Publications, 143–173.

Rodriguez, E., and McDaniel, R. (2001). Combinatorial biosynthesis of antimicrobials and other natural products. *Curr Opin Microbiol* 4:526–534.

Rowan, R., and Powers, D.A. (1992). Ribosomal RNA sequences and the diversity of symbiotic dinoflagellates (zooxanthellae). *Proc Natl Acad Sci USA* 89:3639–3643.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, NY.: Cold Spring Harbor Laboratory Press.

Seow, K.T., Meurer, G., Gerlitz, M., Wendt-Pienkowski, E., Hutchinson, C.R., and Davies, J. (1997). A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J Bacteriol* 179:7360–7368.

SwofFord, D.L. (2000). PAUP* Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, Mass.: Sinauer.

Taylor F.J.R. (1987). General group characteristics; special features of interest; short history of dinoflagellate study. In: *The Biology of Dinoflagellates*, Taylor, F.J.R., ed. Maldent Mass.: Blackwell Scientific Publications, 1–23.

Taylor, F.J.R. (1993). The species problem and its impact on harmful phytoplankton studies. In: *Toxic Phytoplankton Blooms in the sea*, Smayda, T.J., and Shimizu, Y., eds. New York, N.Y.:Elsevier, 81–86.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving sensitivity of progressive multiple sequence alignment through sequence weighing, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.

Tosteson, T.R., Ballintine, D.L., Tosteson, C.G., Hensley, V., and Bardales, A.T. (1989). Associated bacterial flora, growth and toxicity of cultured benthic dinoflagellates *Ostreopsis lenticularis* and *Gambierdiscus toxicus*. *Appl Env Microbiol* 55:137–141.

Zhou, J., and Fritz, L. (1994). Okadaic antibody localizes to chloroplasts in the DSP-toxin-producing dinoflagellates *Prorocentrum lima* and *Prorocentrum maculosum*. *Phycologia* 33:455–461.