Journal of Plant Biology, March 2000, 43(2) : 88-93

Isolation and Sequence Analysis of Promoter DNA Fragments of the Luciferin-Binding Protein Gene in *Gonyaulax polyedra*

Sohyoung Kim, Hee-Gyun Lee, and Dong-Hee Lee*
Department of Biological Science, Ewha Womans University, Seoul 120-750, Korea

To understand the unusual features of the genes and genomes from *Gonyaulax polyedra*, we isolated the promoter portions of the luciferin binding protein (LBP) gene, using IPCR methods, and characterized their sequences. Five LBP genomic clones were classified into a group of genes from the LBPα family, based on the sequence homology of the coding portion of the LBP gene. They were subdivided into two groups. Southern analysis implied that the promoter region is conserved well in most LBP genes. The comparison of the promoter regions from the LBP and luciferase genes showed that, although some portions of their sequences were well conserved, these two genes did not share common features of promoter region, as is normally found in eukaryotes or prokaryotes.

Keywords: dinoflagellates, *Gonyaulax polyedra*, LBP gene, promoter

The dinoflagellates possess a combination of prokaryotic and eukaryotic features, which suggests that they are a geologically old group. They may have diverged from the higher eukaryotic lineage before the evolution of eukaryotic chromatin but after the evolution of repeated DNA. Some of the primitive nuclear features that support a prokaryotic affinity include the fibrillar diameter of dinoflagellate chromatin, a low level of chromosomal basic proteins, membrane attachment of chromosomes, and a swirl pattern observed in sectioned chromosomes. In contrast, such features as a nucleus bound by a porous double membrane; repeated and highly complex DNA; an S-phase of DNA synthesis in the cell cycle; basic proteins; and the reinterpretation of extranuclear microtubules as the spindle support the contention that dinoflagellates are eukaryotes. In addition, although the 5S RNA of the dinoflagellate *Crypthecodinium cohnii* lacks three different regions that are highly conserved in all prokaryotic, but not eukaryotic, 5S RNA molecules, its extensive sequence homology with 5S RNAs from higher eukaryotes and the secondary loop-hairpin structure closely resembles that suggested for other eukaryotic 5S RNAs (Vigne and Jordan, 1977; Hori and Osawa, 1979). This is dissimilar from that predicted for most prokaryotic molecules. Therefore, a separate kingdom, Mesokaryota, has been proposed to accommodate the dinoflagellates (Dodge, 1965).

LBP (luciferin binding protein), which sequesters luciferin, is responsible for light emission with luciferase and luciferin (a tetrapyrrol) (Morse et al., 1989a; Morse et al., 1989b). A large gene family encodes LBP (over 1000 copies per genome). The isolated cDNAs have been classified into one of two groups, LBPα and LBPβ. Members of the same group share at least 98% sequence identity, whereas members of the other group share no more than 86% sequence identity. As part of its organization, the LBP gene contains no introns, but does exhibit certain features not typical of an eukaryotic gene. Its promoter does not include the typical TATA box within ~50 nucleotides upstream of the transcription start site. Likewise, the usual poly (A+) signal (AAUAAA) is not present on the end of the LBP mRNA. However, the universal genetic code and the conserved positions relevant for the translational apparatus are maintained. The goal in studying the LBP genomic genes has been to examine the extent of heterogeneity in the LBP genomic genes and to try to detect any unusual features in the genomic organization and promoter region of LBP, compared with luciferase.

In this current study to understand the previously described characteristics of the genes from *Gonyaulax polyedra*, we first isolated LBP genes using an inverse PCR method, then analyzed their sequences.

MATERIALS AND METHODS

Cells

G. polyedra (GP 70) cells were grown in 2.8-L Fern-

*Corresponding author; fax +82-2-3277-2385
e-mail lee@mm.ewha.ac.kr

bach flasks containing a modified seawater medium (F/2; Guillard and Ryther, 1962), at 20°C under a 16-h light/8-h dark cycle.

Chemicals

The oligonucleotides for inverse PCR were synthesized by either Bioneer (Korea Biotech. Inc.) or the Central Laboratory of College of Pharmacy, Ewha Womans University. A pT7Blue T-Vector kit for the cloning of PCR product was purchased from Novagen. The sequenase version 2.0 DNA sequencing kit was from United States Biochemical; [α-^{35}S] dATP for sequencing and [α-^{32}P]dCTP for the Southern analysis were purchased from Amersham or Dupont NEN. A nylon blotting membrane, Hybond-N, was obtained from Amersham, while the general nucleic acid modifying enzymes were from Promega, NEB, or Amersham. All other chemicals, usually molecular biology grade, were from Sigma.

Preparation of Genomic DNA

The culture, approximately one month old, was harvested via a Nitex filter (20 μm pore size). To isolate the genomic DNA, we followed, with minor modifications, the method described by Rogers and Bendich (1988), utilizing cetyltrimethylammonium bromide (CTAB) to remove the polysaccharides. Frozen cells were ground with liquid nitrogen in a precooled mortar. The powdered material was then transferred to a 50-mL screw-cap centrifuge tube. One volume of hot (65°C) 2x CTAB buffer [(2% (w/v) CTAB, 100 mM Tris-HCl, 20 mM EDTA, 1.4 M NaCl, and 1% polyvinylpyrrolidone (M_r 40,000); pH 8.0] was then added, gently mixed with 1 volume of chloroform/isoamyl alcohol (24:1), and centrifuged at 11,000*g* for 30 s. The top phase then was transferred to a new tube. After adding 0.1 volume of 10% CTAB (10% CTAB and 0.7% NaCl), the chloroform/isoamyl alcohol extraction was repeated.

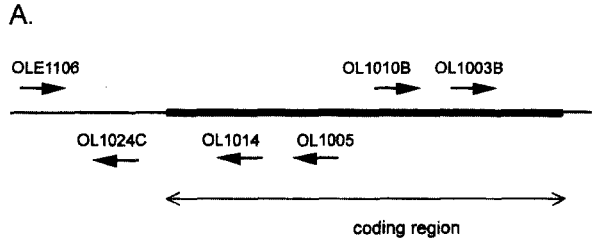
To separate nucleic acids from polysaccharides, 1 volume of CTAB precipitation buffer (10 mM Tris-HCl, 1 mM EDTA, and 1 M NaCl; pH 8.0) was added and then reprecipitated with 2 volumes of 95% ethanol. The nucleic acid was recovered with a hooked glass rod, vacuum-dried, and resuspended in TE buffer. Purity of the DNA was assured by treating the preparation with RNaseA and extracting it successively with phenol, phenol: chloroform, and chloroform, followed by ethanol precipitation (Sambrook et al., 1989). The pellet was resuspended in TE buffer and stored at –20°C.

Cloning of the Genomic LBP Gene by IPCR

For LBP genomic gene cloning and characterization, we designed six oligonucleotides (Fig. 1), based on the published information of the LBP cDNA sequence (Lee et al., 1993). A pair of oligonucleotides, OL1005 and OL1010B, was used as primers for PCR amplification; another pair, OL1014 and OL1003B, was used as primers for our nested PCR amplification. The reaction conditions used for amplification of DNA were 40 cycles with denaturation at 93°C for 1 min, annealing at 61 to 72°C for 30 s, and extension at 72°C for 3 min.

Each reaction, typically in a 20 μL volume, contained 50 mM KCl, 10 mM Tris-HCl (pH 9.0 at 25°C), 0.1% Triton X-100, 1.9 mM $MgCl_2$, 200 mM dNTP, 4 pmole of appropriate primers, and varying amounts of template DNA. The reaction was initiated by adding 1 unit of *Taq* DNA polymerase. Products of the PCR amplification were analyzed on 1% or 1.4% agarose gels.

To prepare the template for IPCR, 1 μg of genomic DNA was restricted with 10 units of EcoRI or PstI at 37°C for 7 to 8 h. This was followed by phenol/chloroform extraction and ethanol precipitation. The DNA was redissolved in 400 μL of ligation buffer [30 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM DTT, and 0.5 mM ATP]. It was then circularized overnight



B.

	5' 3'
OLE 1106	GGAATTCGATTAACTTGGATAA
OL 1024C	ACACGCAACACGTGCACAGCAT
OL 1014	ACTGGTCTTCCACTAACGCATG
OL 1005	AATGAAGTGCTCGAGCCTCAAG
OL 1010B	CCGGCAAAGTTCATGGACGTGT
OL 1003B	ATCAAGCAGCCCTGGAAGGTCA

Figure 1. Primer used for PCR amplification. (A) The position of primers on the LBP gene. (B) Nucleotide sequences of the primers used for PCR.

at 19°C by adding 3 units of T_4 DNA ligase. After ethanol precipitation, the DNA was resuspended in 10 μL TE buffer.

Cloning and Characterization of the PCR Product

We used a pT_7Blue T-Vector kit to clone the PCR product of the LBP genomic gene. The product was ligated, then transformed into *Escherichia coli* NovaBlue competent cells (Novagen). Plasmids from white colonies on the X-gal plates were isolated by alkaline methods; insert size was found through restriction mapping. The sequence of the insert was determined by using a dideoxy chain-termination method, then analyzed via a DNasis program (Hitachi Software); BLAST (Altschul et al., 1990); and the CLUSTA W multiple sequence alignments tool 1.8 program (Thompson et al., 1994; http://dot.imgen.bcm.tmc.edu:9331/multi-align/Options/clustalw.html).

Genomic Southern Blot Analysis

For LBP genomic Southern analysis, the membrane was hybridizaed with ^{32}P-labeled probe of the 5'-portion of the LBP genomic gene (–400 to +5). The 400-bp PCR products, amplified with OL 1024C and

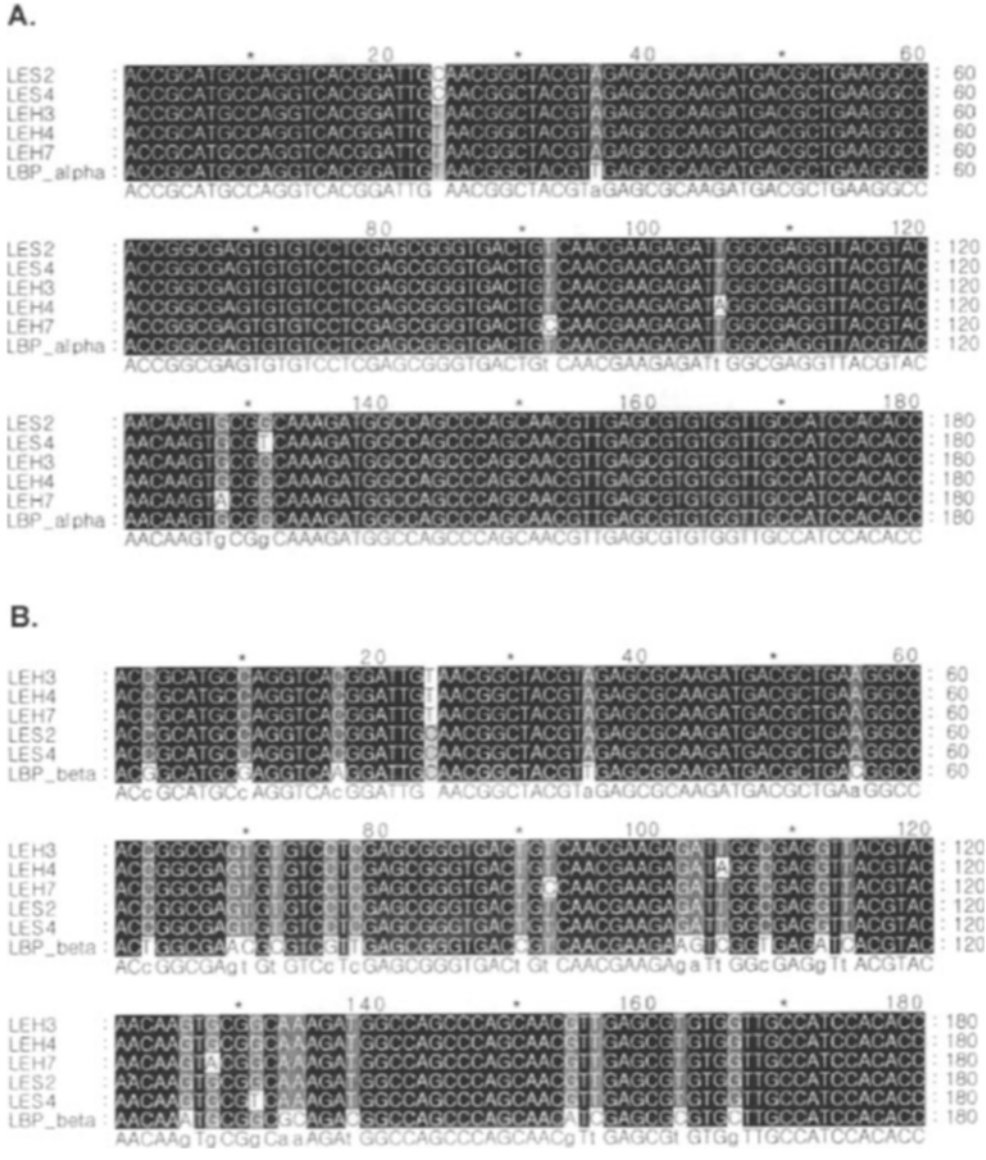


Figure 2. Nucleotide sequence comparison of five different genomic sequences with LBP cDNA sequences. (A) Comparison with LBPα, (B) Comparison with LBPβ. Box indicates sequence variations.

OLE 1106, were isolated and radioactively labeled using [α-^{32}P]dCTP (3000 Ci/mmol) with a Prime-It II random primer labeling kit (Stratagen). The radioactively labeled probe was denatured at 80°C for 10 min and used for hybridization.

The membrane was prehybridized at 42°C for 30 min, and hybridized with radiolabelled probes at 42°C overnight. It was then washed twice in 2x SSC; 0.5% SDS at room temperature for 30 min; and twice in 0.5x SSC, 0.5% SDS at 65°C for 15 min. The filters were exposed for 3 days to X-ray film.

RESULTS AND DISCUSSION

Isolation and Sequence Analysis of LBP Genomic Clones

After obtaining five genomic clones, we used the Clustal W tool to compare their putative coding portions with the published LBP cDNA sequences (Lee et al., 1993; Machabee et al., 1994). The cloned gene fragments (Fig. 2) had higher sequence homology with the LBPα group (>97%) than with the LBPβ group (84 to 86%). Therefore, we concluded that the isolated genomic clones were part of the active LBPα genes.

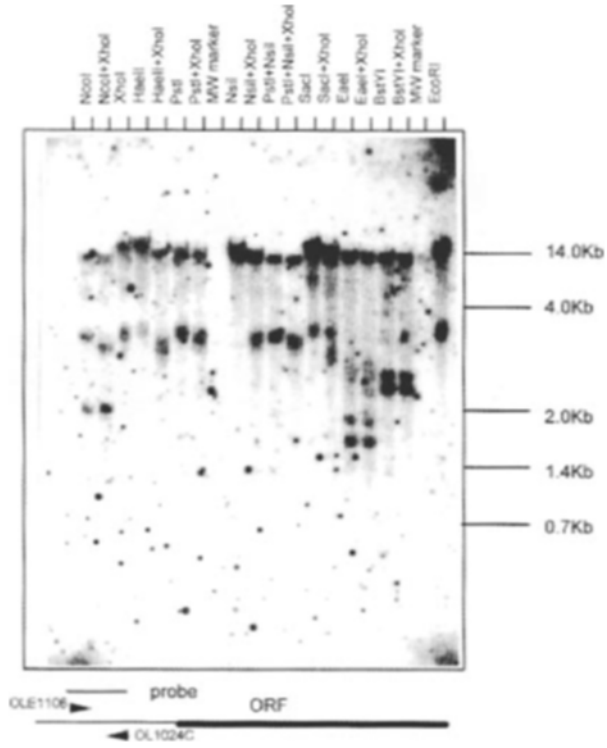


Figure 3. Southern blot analysis of the LBP promoter region. Hybridization was carried out with a labeled probe of the 5' portion (+4~−400) of the LBP genomic clone at 42°C for 12 h. The filter was washed and exposed to X-ray film with an intensifying screen at −70°C for 3 days.

Genomic Southern Blot Analysis of the LBP Promoter Region

Southern blot analysis with our 400-bp ^{32}P-labeled probe showed that one or two strongly hybridized signals existed (Fig. 3). This suggests that the sequences, especially on the recognition sites of our most restriction enzymes, were well conserved. We concluded that the promoter portion was well conserved in most of the genes in the LBPα group.

DNA Sequence Analysis of the 5' Portion

Following multiple alignments of the 5' portions of the DNA sequences, our five clones showed >97% sequence identity in the −36 ~ −1 region (Fig. 4). (Nucleotide number −1 was assigned based on the published transcription start site of the LBP genes; Lee et al., 1993). Much less homology was found at −100 ~ −36, which suggests that the five clones were part of two separate groups, i.e., "group H" (LES2, LES4, and LES7) and "group S" (LEH3 and LEH8). Therefore, more than two different groups of LBP genomic genes can generate LBPα mRNA. This observation confirmed our hypothesis during cloning about the variability of the EcoRI site on the upstream region.

Sequence Comparison of Promoter Portions of the LBP and Luciferase Genes

For *G. polyedra*, only three genomic clones have been published: the LBP gene, the luciferase gene (Li and Hastings, 1998), and the peridinin-chlorophyll a-binding protein gene (Le et al., 1997). In the current study, we characterized the conserved sequence among three promoter regions in *G. polyedra,* the 5' region of two clones (LES2 and LEH3), and two published genomic clones, LBP and luciferase as shown in Figure 5. A novel sequence, CGTGAACGCAGTG, is present at around −540 upstream from the transcription site of the peridinin-chlorophyll a-binding protein gene (Le et al., 1997). We found that same sequence at around −115 upstream from the transcription site for the luciferase gene, but not for the LBP gene.

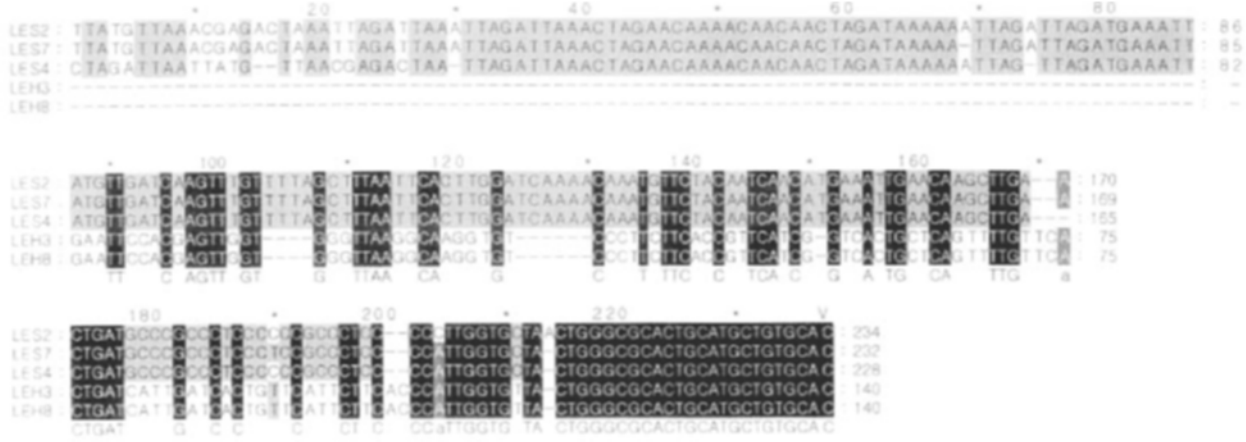


Figure 4. Comparison of nucleotide sequences of promoter portions of LBP clones with the previously published genomic LBP sequence, pDL. Conserved promoter regions are boxed. The transcription start sites are marked by an arrowhead.

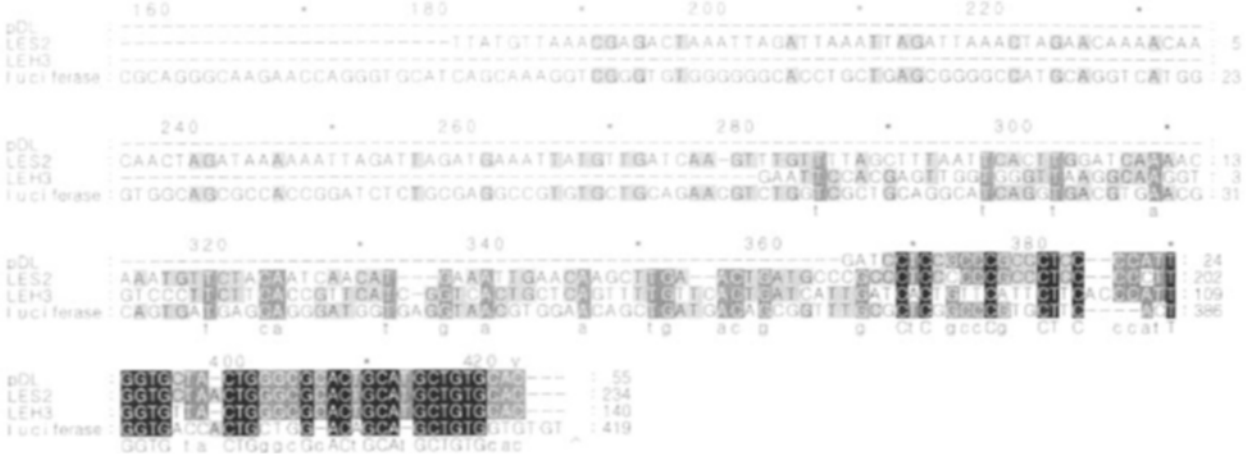


Figure 5. Nucleotide sequence comparison of the promoter regions of the LBP and luciferase genes. The promoter regions are aligned from the transcription start point by Clustal W multiple alignment programs. The transcription start points of the LBP and luciferase genes are marked by arrowheads. Box indicates conserved nucleotide sequences.

In addition, rather than the typical TATA box-like sequence in the upstream sequences, the 5' regions of all three genes had high GC ratios and several GC patches at positions −54 ~ −3. Other studies have shown that some of the eukaryotic promoters lack such a TATA box. For example, the promoters of *c-rel, nfkb1, nfkb2,* and *Ikb1* are expressed at their highest levels in cells of lymphoid origin, in a GC-rich region (kB sites); these may serve instead (Liptay et al., 1994). Likewise, we found that all three *Gonyaulax* promoter regions did contain such GC-rich sequences in the regions from positions −54 ~ −3, which might functionally replace the TATA box.

We conclude from this study that *G. polyedra* could have a different mechanism of transcription control compared with other eukaryotic organisms. This species may not share common transcription factors, and would have unusual GC-rich regions, even in the conserved potential promoter sites. This unique feature implies that *G. polyedra* has its own typical conserved sequence, probably for RNA polymerase binding or transcription factor binding. That mechanism for transcription machinery, therefore, would differ from the known mechanism in both eukaryotes and prokaryotes.

ACKNOWLEDGMENT

We thank Prof. Hastings and Dr. Li of Harvard Uni-

versity for providing sequence information for luciferase. This research was partly supported by a grant for BSRI of the National R&D Project, and by a grant women university of the National R&D Project.

Received April 3, 2000; accepted May 29, 2000.

LITERATURE CITED

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410

Dodge JD (1965) Chromosome structure in the dinoflagellates and the problem of the mesokaryotic cell. Excerpta Med Int Congr Ser 91: 339-341

Guillard RRL, Ryther JH (1962) Studies of marine planktonic diatoms I. *Cyclosella mana* Hustedt and *Detonula confervacea* (Cleve) Gran. Canad. J Microbiol 8: 229-239

Hori H, Osawa S (1979) Evolutionary change in 5S RNA secondary structure and a phylogenic tree of 54 5S RNA species. Proc Natl Acad Sci USA 76: 381-385

Le QH, Markovic P, Hastings JW, Jovine RV, Morse D (1997) Structure and organization of the peridinin-chlorophyll a-binding protein gene in *Gonyaulax polyedra*. Mol Gen Genet 255: 595-604

Lee DH, Mittag M, Sczekan S, Morse D, Hastings JW (1993) Molecular cloning and genomic organization of a gene for Luciferin-binding protein from the Dinoflagellate *Gonyaulax polyedra*. J Biol Chem 268: 8842-8850

Li L, Hastings JW (1998) The structure and organization of the luciferase gene in the photosynthetic dinoflagellate *Gonyaulax polyedra*. Plant Mol Biol 36: 275-84

Liptay S, Schmid RM, Nabel EG, Nabel GJ (1994) Transcriptional regulation of NF-kappa B2: evidence for kappa B-mediated positive and negative autoregulation. Mol Cell Biol 14: 7695-703

Machabee S, Wall L, Morse D (1994) Expression and genomic organization of a dinoflagellate gene family. Plant Mol Biol 25: 23-31

Morse D, Frits L, Hastings JW (1989a) Role of a luciferin-binding protein in the circadian bioluminescent reaction of *Gonyaulax polyedra*. J Biol Chem 264: 11822-11826

Morse D, Milos PM, Roux E, Hastings JW (1989b) Circadian regulation of bioluminescence in *Gonyaulax* involve translational control. Proc Natl Acad Sci USA 86: 172-176

Rogers SO, Bendich AJ (1988) Extraction of DNA from plant tissues, *In* SB Gelvin, RF Schilperoort, DS Verma, eds, Plant Molecular Biology Manual, Vol I. Kluwer Academic Publisher, Belgium, A6: 1-10

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning, A Laboratory Manual, 2nd ed, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680

Vigne R, Jordan BR (1977) Partial enzyme digestion studies on *Escherichia coli, Pseudomonas, Chlorella, Drosophila,* HeLa and yeast 5S RNAs support a general class of 5S RNA models. J Mol Evol 10: 77-86