Hydrobiologia (2009) 616:229–239
DOI 10.1007/s10750-008-9596-5

JELLYFISH BLOOMS

Comparative analysis of nuclear ribosomal DNA from the moon jelly *Aurelia* sp.1 (Cnidaria: Scyphozoa) with characterizations of the 18S, 28S genes, and the intergenic spacer (IGS)

Jang-Seu Ki · Il-Chan Kim · Jae-Seong Lee

Published online: 21 October 2008

Abstract Nuclear ribosomal DNAs (rDNA) constitute a multi-gene family with tandemly arranged units linked by an intergenic spacer (IGS). Here we present the complete DNA sequence (7,731 bp) of a single repeat unit of an rDNA sequence from the moon jelly *Aurelia* sp.1 (Cnidaria: Scypozoa). The tandemly repeated rDNA units consisted of coding and non-coding regions, whose arrangement was 18S rDNA (1,814 bp, 46.2% of GC content)-internal transcribed spacer 1 (ITS1: 272 bp, 39.7%)-5.8S rDNA (158 bp; 50.7%)-ITS2 (278 bp, 51.4%)-28S rDNA (3,606 bp, 49.7%)-IGS (1,603 bp, 45.6%). GC composition in the single unit of rDNA was 47.8%. None of the 5S rDNA was found in the repeat units. Putative structures of a termination transcription signal (poly(T) tract) and promoter-like bi-repeats within the non-coding region were also identified. A block of minisatellites with five repeats was detected within the IGS. Comparative analyses of parsimony and dot plots showed that the IGS was highly informative. The sequence revealed here was the first completion of rDNA from the phylum Cnidaria, using as a model of rDNA for making molecular comparisons of jellyfish members.

Keywords Jellyfish · Nuclear rDNA · Transcription repeat unit · Minisatellite

Guest editors: K. A. Pitt & J. E. Purcell
Jellyfish Blooms: Causes, Consequences, and Recent Advances

J.-S. Ki
Department of Molecular and Environmental Bioscience, Graduate School, Hanyang University, Seoul 133-791, South Korea

I.-C. Kim
Polar BioCenter, Korea Polar Research Institute, Korea Ocean Research and Development Institute, Incheon 406-840, South Korea

J.-S. Lee (✉)
Department of Chemistry, and the National Research Lab of Marine Molecular and Environmental Bioscience, College of Natural Sciences, Hanyang University, Seoul, 133-791, South Korea
e-mail: jslee2@hanyang.ac.kr

Introduction

The moon jelly *Aurelia* Péron & Lesueur, 1810 (Cnidaria; Scyphozoa), is one of the most common and widely distributed species of jellyfish (Arai, 1997). At least 12 species of *Aurelia* have been described based on morphological differences of the medusae (Mayer, 1910; Kramp, 1961). At present, however, only three species, *A. aurita* Linnaeus, 1758, *A. labiata* Chamisso & Eysenhardt, 1821, and *A. limbata* Brandt, 1835, are accepted by systematists. Recent molecular tools, particularly DNA sequencing, can discriminate the species more clearly, allowing taxonomic revisions and molecular phylogenetic inference studies of the jellyfish (Schroth et al., 2002; Dawson et al., 2005; Collins et al., 2006).

For the moon jelly *Aurelia*, Dawson and colleagues (Dawson & Jacobs, 2001; Dawson & Martin, 2001; Dawson, 2003; Dawson et al., 2005) have constructed a more acceptable taxonomic system by combining their morphological characteristics and DNA sequences. Genetic information is generally accepted as a useful molecular tool to discriminate between jellyfish species.

The ribosomal DNA (rDNA) is the region of the genome coding for the RNA component of ribosomes. Eukaryotic nuclear rDNA is tandemly organized, with copy numbers up to the order of 10,000 (Schlötterer, 1998). Each repeat unit consists of the genes coding for the 18S, 28S, and the 5.8S rDNA. These coding regions are separated from each other in the primary transcript by the internal transcribed spacer (ITS) as well as by intergenic spacer (IGS). The rDNA coding regions have remained relatively constant within the same taxa, making these DNA regions the important information sources for the study of phylogenetic relationships. The ITS and IGS rDNA as non-coding regions can be useful to reconstruct relatively recent evolutionary events (Hillis & Dixon, 1991). Thus, complete rDNA sequences provide us with various options depending on the variability of the molecules. By searching all databases, rDNAs from jellyfish have been partially sequenced in a region spanning the 18S and partial 28S (Dawson et al., 2005). None of the complete structure on the transcription repeating unit of rDNA from Cnidarian members, including jellyfish, has been reported.

In Korean coastal environments, the moon jelly *Aurelia* was recently found and its dense blooms caused economic losses for fisheries and power plants (data from the National Fisheries Research and Development Institute, http://nfrdi.re.kr/). Recently, Ki et al. (2008) confirmed that all the moon jellies blooming in different areas of Korea had an identical genotype (*Aurelia* sp.1) using mitochondrial COI gene and nuclear ITS-5.8S rDNA sequences. Here, we present the entire nucleotide sequence of a single unit of rDNA from Korean *Aurelia* sp.1 and a characterization of the rDNA. Comparative analyses of parsimony and dot plot were performed with some known complete rDNA sequences in order to better understand the rDNA relationships among other eukaryotes.

Materials and methods

Sample collection

Five specimens of *Aurelia* sp.1 were collected from different localities in western Korean coastal waters (Incheon; 37°26′23″N, 126°22′40″E). After sampling, the individuals were immediately transferred into absolute ethanol to dehydrate and stored at room temperature until use. Genomic DNA was isolated using a previously described procedure (Lee, 2000). The DNA was purified with the DNeasy tissue kit (Qiagen, Valencia, CA) according to the manufacturer's instructions.

Long and typical polymerase chain reaction

Polymerase chain reaction (PCR) primers (Table 1) were designed for amplification of the entire rDNA from *Aurelia*, based on a comparison of several eukaryotic rDNA sequences. Specifically, we developed primers targeting 28S rDNAs, by comparing sequences of *Atolla vanhoeffeni* Russell, 1957 (GenBank accession no. AY026368), *Catostylus* sp. Agassiz, 1862 (AY920777), *Chrysaora melanaster* Brandt, 1835 (AY920780), *Craterolophus convolvulus* Johnson, 1835 (AY920781), *Haliclystus sanjuanensis* Hyman, 1940 (AY920782), *Nausithoe rubra* Vanhöffen, 1902 (AY920776), and *Phacellophora camtschatica* Brandt, 1838 (AY920778).

Long and accurate (LA)-PCR amplification was carried out from purified genomic DNA and three primer pairs (e.g., JF-18F24, JF-28R1; JF-28F0, JF-28R3.3K; JF-18R70, JF-28F3.2K). PCR reactions were carried out in 1× PCR buffer (10 mM Tris–HCl, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin; pH 8.3) with <0.1 μg genomic DNA template, 200 μM each of the four dNTPs, 0.5 μM of each primer and 0.2 units of LA *Taq* polymerase (TaKaRa, Japan). Using a Thermoblock (iCycler, Bio-Rad, CA), PCR thermocycling parameters were as follows: 95°C for 3 min; 35 cycles of denaturation at 95°C for 20 s, annealing at 55°C for 30 s and extension at 68°C for 5 min, and a final extension at 72°C for 10 min. The PCR products (2 μl) were analyzed by 1.0% agarose gel electrophoresis according to a standard method.

Table 1 Primers used for PCR and sequencing (Seq.) of complete rDNA unit from *Aurelia* sp.1

Primer	Nucleotide sequence (5′–3′)	Location[b]	Application(s)
JF-18F24[a]	TGGTTGATCCTGCCAGTAG	18S (5–23)	PCR/Seq.
JF-18R70	CGCAGTTTCACAGTACAAGTGC	18S (90–69)	PCR/Seq.
JF-28F3.2K	AGGGAACGTGAGCTGGGTTTAG	28S (3155–3176)	PCR/Seq.
JF-28R1	ACGCTTCTCGAGACTACAATTCGC	28S (164–141)	PCR/Seq.
JF-28R3.3K	ATCTGCGGTTCCTCTCGTAC	28S (3261–3242)	PCR/Seq.
JF-28F0	AAGGATTCCCTCAGTAACGG	28S (72–91)	PCR/Seq.
MJ-18F900	TTCTTGGATTTACGAAAGAC	18S (921–940)	Seq.
MJ-18F1778	CCGAGAAGTCGCTCTAGTTC	18S (1730–1749)	Seq.
JF-IGSR1	GACTACTGGCAGGATCAACC	18S (25–6)	Seq.
JF-28F1	AGTCGGGTTGCTTGGGAATGCAGC	28S (269–292)	Seq.
JF-28F2	CGATAGCGAACAAGTACCGTGAG	28S (341–363)	Seq.
JF-28F1390	GAACCGAACGCTGAGTTAAG	28S (968–986)	Seq.
MJ-IGS-F1	GTTGCATTGCGGCTGAGTG	IGS (128–146)	Seq.
MJ-IGS-F900	TTAGGATACCGAAATCAC	IGS (513–530)	Seq.
MJ-IGS-F1390	TGGCTCGAGGCTGACATTG	IGS (1356–1375)	Seq.
MJ-IGS-R1k	AAAGTGAACCTGGCAGAC	IGS (680–663)	Seq.

[a] F, forward primer, R, reverse primer, S, sequencing primer

[b] Locations refer to the *Aurelia* rDNA numbering (GenBank No. EU276014) revealed here

For direct DNA sequencing, PCR amplicons were purified with QIAquick PCR purification Kit (Qiagen GmbH, Germany). DNA sequencing reactions were performed in a ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (P/N 4303149, PE Biosystems, CA) using the PCR products (2 µl) as the template and 10 picomoles of sequencing primers (Table 1). Labeled DNA fragments were analyzed on an automated DNA sequencer (Model 3700, Applied Biosystems, CA).

Gene identification and secondary structures

Editing and contig assembly of the rDNA sequence fragments were carried out with Sequencher 4.7 (Gene Codes, MI, USA). The coding rDNA sequences were identified by considering those of other eukaryotes in the NCBI database and typical nucleotide sequences such as "TAT CTG G" for the start of 18S rDNA, "TTT GT" for the end of 28S rDNA (see Ki & Han, 2007). The sequence determined here has been deposited to GenBank with the accession number EU276014.

Putative secondary structures of assumed termination and promoter signals within the rDNA IGS were estimated using the program Mfold, version 3.2 (http://www.bioinfo.rpi.edu/applications/mfold/old/rna/) according to Zuker (2003). With the default option (e.g., temperature setting, $T = 37°C$), Mfold predicted six secondary structures from the entire IGS sequence. Different parameter settings of, for example, temperature ($T = 10°C$, $T = 20°C$, $T = 37°C$) did not affect the general architecture, but did result in different energy levels for the secondary structures. Among them, we finally selected a general model of secondary structures of termination and promoter-like signals. Secondary structure models inferred here were redrawn with RNA structure ver. 4.5 (Mathews et al., 1999).

Data analysis

General molecular features of the *Aurelia* rDNA were calculated by Genetyx ver.7.0 (Hitachi Engineering Co. Ltd., Japan) and MEGA 4.0 (Kumar et al., 2001). In addition, nucleic acid distribution, sequence complexity, and entropy across the entire rDNA nucleotides of *Aurelia* sp.1 were calculated with the BioAnnotator in Vector NTI Advance 10.3.0 (Invitrogen). Repeat sequence patterns in the rDNA IGS sequences were analyzed using the Genetyx 7.0 and Tandem Repeats Finder (http://tandem.bu.edu/trf/trf.basic.submit.html).

For comparative molecular features, the rDNA of *Aurelia* sp.1 was compared with four complete rDNAs from *Meloidogyne artiellia* Franklin, 1961 (Nematoda, AF248477), *Anopheles albimanus* Wiedemann, 1820 (Arthropoda, L78065), *Chironomus tentans* Fabricius, 1805 (Arthropoda, X99212), and *Herdmania momus* Savigny, 1816 (Ascidiacea, X53538), which were retrieved from DDBJ/EMBL/GenBank. For parsimony analysis, sequence alignment was performed with the five complete rDNAs, using Clustal W ver. 1.4 with the default settings for gap inclusion and extension. Various regions were further aligned manually using the BioEdit 5.09 (North Carolina State University, NC). Genetic distance values were calculated by using the aligned DNA sequences according to the Kimura 2-parameter model (Kimura, 1980), and molecular similarity was measured in BioEdit 5.0.6. Further, comparative analyses such as pairwise (p) distance and parsimony-informative (P-I) sites, transition/transversion ratio were implicated with MEGA 4.0 with the above data matrix. Dot-plot analysis was carried out using the MegAlign 5.01 software (DNAstar Inc., WI).

Results

General features of complete rDNA unit of Aurelia sp.1

New PCR primers ($N = 6$) were designed for the isolation of the entire rDNA of *Aurelia* sp.1 (Table 1). LA-PCRs, by employing three primer pairs selected for amplification of less than 4 kb, successfully amplified the expected sizes of PCR fragments from genomic DNA of *Aurelia* sp.1 (data not shown). Using the DNA fragments obtained by PCR amplifications, we could carry out complete sequencing of the entire rDNA sequence of *Aurelia* (Fig. 1). By BLAST search, partial DNA sequences obtained from 18S- and 28S-containing PCR amplicons were matched to some sequences revealed from Cnidaria, including *Aurelia*. We observed that the 18S sequence matched well with those of *A. aurita* (AY039208), *Aurelia* sp. (AY920770), *Chrysaora melanaster* (AF358099), and *Stomolophus meleagris* Agassiz, 1862 (AF358101), and the 28S sequence matched those of *Phacellophora camtschatica* Brandt, 1835 (AY920778), and *Chrysaora melanaster* (AY920780). This confirmed that the new PCR primers successfully amplified the expected rDNA from *Aurelia*.

A total length of single rDNA repeat unit from *Aurelia* sp.1 was determined to be 7,731 bp. It was organized in the typical eukaryotic fashion of rDNA, i.e. 18S-ITS1-5.8S-ITS2-28S-IGS (Fig. 1). Specifically, DNA sequence of *Aurelia* rDNA was recorded at 1,814 bp (18S rDNA), 272 bp (ITS1), 158 bp (5.8S rDNA), 278 bp (ITS2), 3,606 bp (28S rDNA), and 1,603 bp (IGS). Each rDNA transcription unit was separated by non-coding regions (i.e. IGS). The IGS was assumed to comprise 3′ external transcribed spacer (ETS), 5′ ETS, and non-transcribed spacer (NTS). Intron-like sequences were not detected in 18S and 28S rDNA coding regions.

Nucleotide frequencies of complete rDNA were recorded at A, 26.1%; T, 26.1%; G, 26.8%; C, 21.0% (Table 2). Among them, cytosine (C) content was the rarest and GC content was 47.8% (AT, 52.3%). Specifically, base frequencies of each rDNA locus were identical. In addition, each nucleotide composition was nearly identical among the rDNA coding and non-coding genes excluding ITS1. Base composition of ITS was recorded at A, 27.6%; T, 32.7%; G, 22.4%; and C, 17.3%.

Characteristics of Aurelia *rDNA IGS*

The IGS region of *Aurelia* contained 1,603 nucleotides and its GC content was 46.2%. In some invertebrates (e.g., *Calanus finmarchicus* Gunner, 1765, X06056), 5S rDNA as the smallest subunit of the rRNA coding genes is located within the 18S–28S rDNA repeats (Drouin et al., 1987). Searching the 5S rDNA database (http://rose.man.poznan.pl/5SData/), no pattern of 5S rDNA sequences in IGS of *Aurelia* was found. The IGS consisted of putative transcription termination and bi-repeated sequences. We detected a poly(T) tract (TTT ATT TTT ATT) in 5′ ETS region adjacent to the end of 28S rDNAs, and AGG CCG T(A)G T(G)G T(G)G in 3′ ETS region located upstream of the 18S rDNA (Fig. 1B). As predicted by structural models, the Mfold analysis showed that nucleotides between 28S and the termination signal sequences formed a stem and loop structure (Fig. 2A). On the other hand, some nucleotides of the bi-repeat sequences formed a hair-pin structure (Fig. 2B). Interestingly, a block of

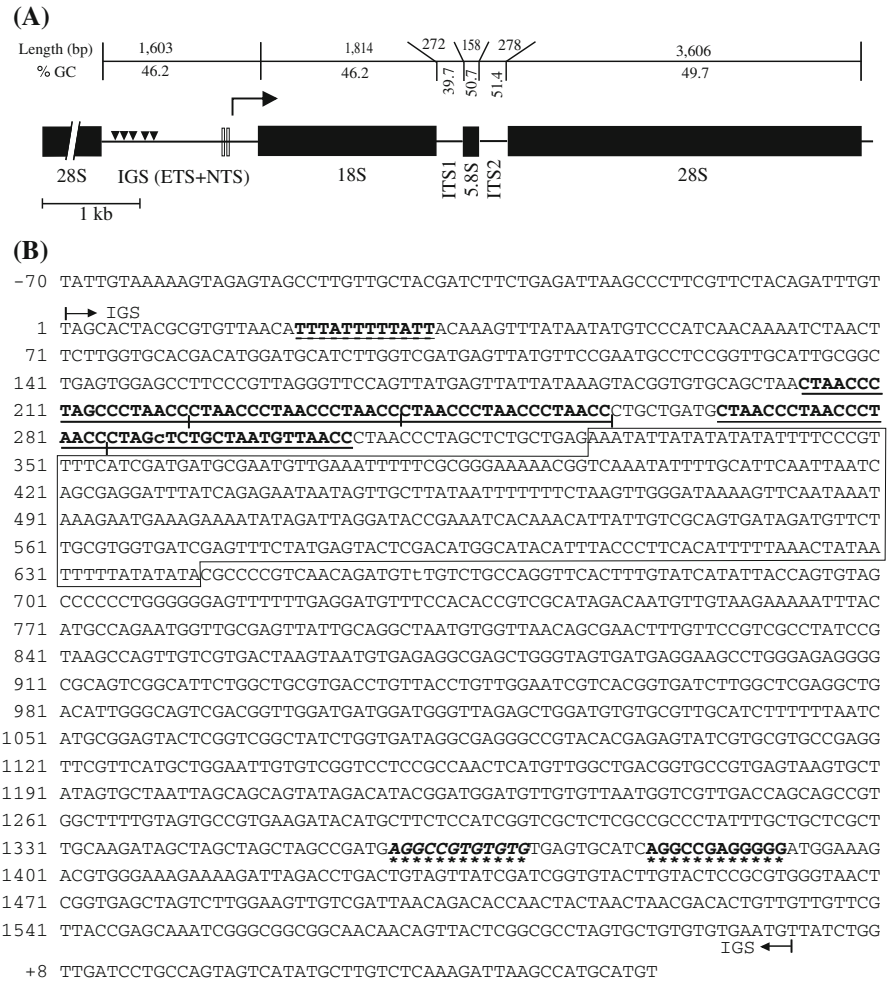

Fig. 1 Schematic presentation of the single unit of rDNA complex (**A**) and nucleotide sequence (**B**) of *Aurelia* sp. Solid boxes indicate the ribosomal RNA genes, and thin lines represent ITS or IGS. Nucleotide sequences in length and GC composition of each locus are represented on/under line by calculation from a single unit of rDNA. The putative transcription start site is represented by an arrow; solid inverted-triangles represent sub-repeats in IGS. In IGS sequence (**B**), bi-repeats are indicated in asterisks, and minisatellite-like nucleotides are marked in lines. A putative termination signal (poly(T) tract) is represented under a dot line. A box in IGS sequence represents a AT-rich region

Table 2 Nucleotide composition and length of each rDNA locus from *Aurelia* sp.1

Locus	Nucleotide composition (%)				Length (bp)
	A	T	G	C	
18S	26.5	27.3	26.2	20	1,814
ITS1	27.6	32.7	22.4	17.3	272
58S	24.1	25.3	26.6	24.1	158
ITS2	25.2	23.4	25.5	25.9	278
28S	26.4	23.9	28.2	21.5	3,606
IGS	24.9	29.4	25.1	20.5	1,603
Total	26.1	26.2	26.8	21	7,731

minisatellite, "CTA ACC CTA GCC CTA ACC"-like nucleotide sequences with five sub-repeats, was detected between the transcription termination signal and the bi-repeat sequences (Fig. 2).

Nucleic acid distribution and sequence complexity

Taking into account the new genetic information presented here, we constructed a map of rDNA organization and calculated their molecular characteristics, including GC content, thiamine (T) distribution, sequence complexity, and entropy (Fig. 3). Maps of characteristics along the entire rDNA were obtained using sliding windows of 100 nucleotides. The distribution of GC content varied around 50% across the complete rDNA. However, some zones of both ITS and IGS showed considerably low GC content. In contrast, the T distribution fluctuated. We found that the low GC content was caused by a high content of nucleotide T within the IGS. In addition to this, sequence variability was analyzed with sequence complexity and entropy

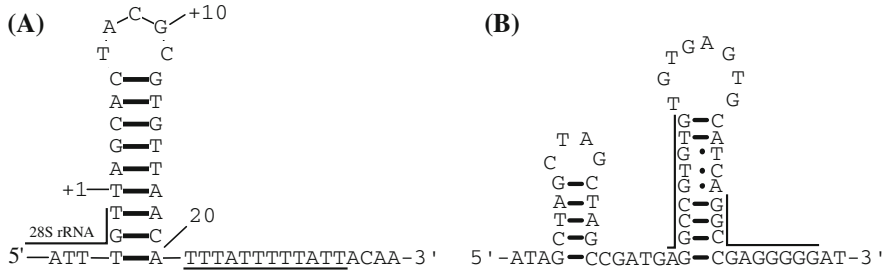

Fig. 2 Putative secondary structure models for termination-associated sequence (**A**) and bi-repeats (**B**), which are marked in lines, within the IGS of *Aurelia* sp.1

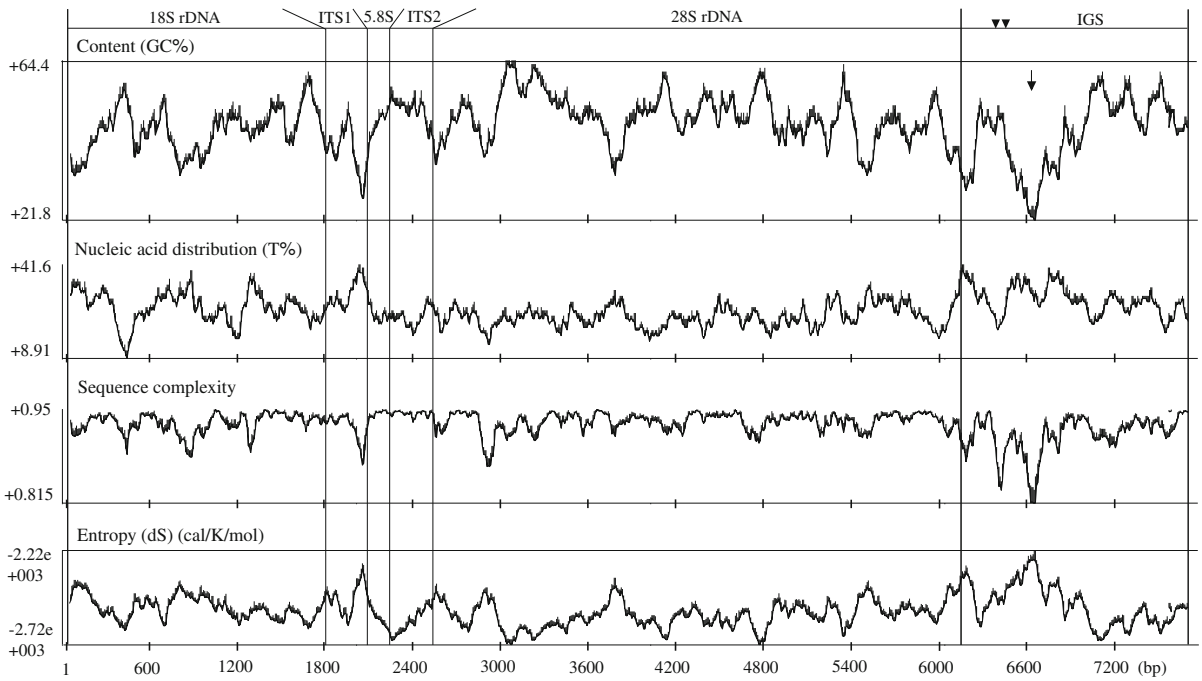

Fig. 3 GC content (%), nucleic acid distribution (% thiamine), sequence complexity and entropy (dS) in 100-bp windows across the entire rDNA nucleotides of *Aurelia* sp.1. Reverse triangles represent the locations of minisatellites within the IGS rDNA. A downward arrow represents a termination transcription signal of poly(T) tract

plotting. Overall, the two variables fluctuated against one another along the rDNA. These observations show a clear difference in profiles of the coding and other non-coding regions. The sequence complexity was considerably higher in non-coding regions such as ITS and IGS, than the coding region. Notably, a zone recording the lowest complexity corresponded to the internal subrepeats and poly(T) regions within the IGS.

Comparative analyses of entire rDNA

The rDNA sequence of *Aurelia* was compared with those of the other eukaryotic members, including *Meloidogyne artiellia* (Nematoda), *Anopheles albimanus* (Arthropoda), *Chironomus tentans* (Arthropoda), and *Herdmania momus* (Ascidiacea; Table 3). Comparative analysis showed that the transition:transversion ratio (Ts/Tv) was higher in coding genes than in non-coding genes; 1.039 in 18S rDNA, 0.951 in 28S, 0.84 in ITS and 0.809 in IGS. High *p*-distance, in contrast, was recoded at ITS (1.710) and IGS (1.509) when compared with those in 18S and 28S. In addition, parsimony analyses showed that the non-coding genes of rDNA contained most P–I sites; 15.1%, 18S; 27.9%, ITS and 19.5%, 28S; 29.9%, IGS. Within coding genes, the P–I of 28S rDNA was higher than that of 18S, probably due to hypervariable regions within the 28S (Hassouna et al., 1984). Variation of non-coding rDNAs was ~1.5 times greater than that of the coding genes as judged by the comparison of P–I sites.

Table 3 Sequence characteristics of rDNA locus among five complete rDNAs from *Aurelia* sp.1 (Cnidaria, accession No. EU276014), *Meloidogyne artiellia* (Nematoda, AF248477), *Anopheles albimanus* (Arthropoda, L78065) *Chironomus tentans* (Arthropoda, X99212), and *Herdmania momus* (Ascidiacea, X53538), respectively. *p*-distances were calculated with the Kimura 2-parameter model

Locus	Nn	Nc	Nv	Ts	Tv	Ts/Tv	P-I	%P-I	*p*-distance
18S rDNA	2044	1038	919	242	233	1.039	308	15.1	0.324
ITS, 5.8S	832	130	661	166	187	0.888	228	27.4	1.097
ITS1, 2	664	63	567	137	163	0.84	185	27.9	1.710
28S rDNA	4228	1838	2179	557	586	0.951	826	19.5	0.403
IGS	2245	203	1735	445	550	0.809	671	29.9	1.509

Note: Nn, total number of sites; Nc, total number of conserved sites; Nv, total number of variable sites; Ts, transition; Tv, transversion; P-I, parsimony-informative site

In order to determine whether there was any sequence homology between parallel sequence alignments, the rDNA of *Aurelia* was compared with those of the above eukaryotes by dot-matrix analysis (Fig. 4). The dot plots were obtained using sliding windows of 60 nucleotides along the entire rDNAs. The dots and lines in Fig. 4 represent regions of homology between sequence pairs. Overall, high similarities of the complete rDNAs between *Aurelia* and Nematoda, Insecta and Chordata were recorded at coding regions (e.g., 18S, 5.8S, 28S) rather than the other non-coding regions. No homologous sequence zones within non-coding rDNAs were detected.

Discussion

In this study, the complete rDNA mapping of *Aurelia* sp.1 was achieved for the first time. The gene arrangement of *Aurelia* sp.1 was identical to those of the typical eukaryote rDNA in order (e.g., 18S-5.8S-28S rDNA, no 5S rDNA). Also, none of the insertion or deletion sites were found within the coding regions, when compared with rDNAs from other eukaryotes such as *Meloidogyne artiellia* (Nematoda, AF248477), *Anopheles albimanus* (Arthropoda, L78065) *Chironomus tentans* (Arthropoda, X99212), and *Herdmania momus* (Ascidiacea, X53538).

The 18S rDNA sequence of *Aurelia* contained 1,814 nucleotides, which was generally similar in size to that of other Cnidaria 18S rDNAs (e.g., *Antipathes galapagensis,* Anthozoa, AF100943, 1,815 bp and *A. aurita,* Scyphozoa, AY039208, 1,808 bp). However, it was shorter than that of *Atolla vanhoeffeni* (Anthozoa, AF100942, 1,825 bp) and longer than those of most Cnidaria such as *Cassiopea* sp. (Scyphozoa, AF099675, 1,802 bp), *Tripedalia cystophora* Conant, 1897 (Cubozoa, L10829, 1,802 bp), *Corynactis* sp. Allman, 1846 (Anthozoa, AJ133559; 1,798 bp), *Coryne pusilla* Gaertner, 1774 (Hydrozoa, AJ133558, 1,797 bp), *Anemonia sulcata* Pennant, 1777 (Anthozoa, X53498, 1,799 bp), *Tubastraea coccinea* Lesson, 1829 (Anthozoa, AJ133556, 1,797 bp), and *Parazoanthus axinellae* Schmidt, 1862 (Anthozoa, U42453, 1,795 bp). On the other hand, GC contents of the 18S rDNA sequences were nearly identical around 46% among the sequences of Cnidarian members.

So far, many sequences of *Aurelia* rRNA have been deposited in the public databases (e.g., EMBL/DDBJ/GenBank). However, most of them are either partial or complete sequences of individual rDNA sequences such as 18S, ITS-5.8S, and 28S (Dawson et al., 2005). None of the complete regions from the 18S to 28S rDNA loci has been sequenced as yet. Therefore, we could not compare sequences at different loci, due to intraspecific variations. The present nuclear rDNA sequence, composed of the 18S, ITS, and 28S rDNA sequences, can be used as a reference sequence to compare the different sequences of loci. By BLAST search, we could compare *Aurelia* sp.1 to other *Aurelia* genotypes from various geographical regions. A comparison of 18S rDNA 1,765 sites showed that DNA similarities between the current *Aurelia* sp.1 were 99.2% with *A. aurita* (AY039208), 99.6% with *A. aurita* (AY428815), and at 98.9% with *Aurelia* sp.2 (AY920770), suggesting that the 18S rDNA may be highly conserved within the genus *Aurelia*. In contrast, ITS comparisons showed that DNA similarities were quite different. For example, highest similarity

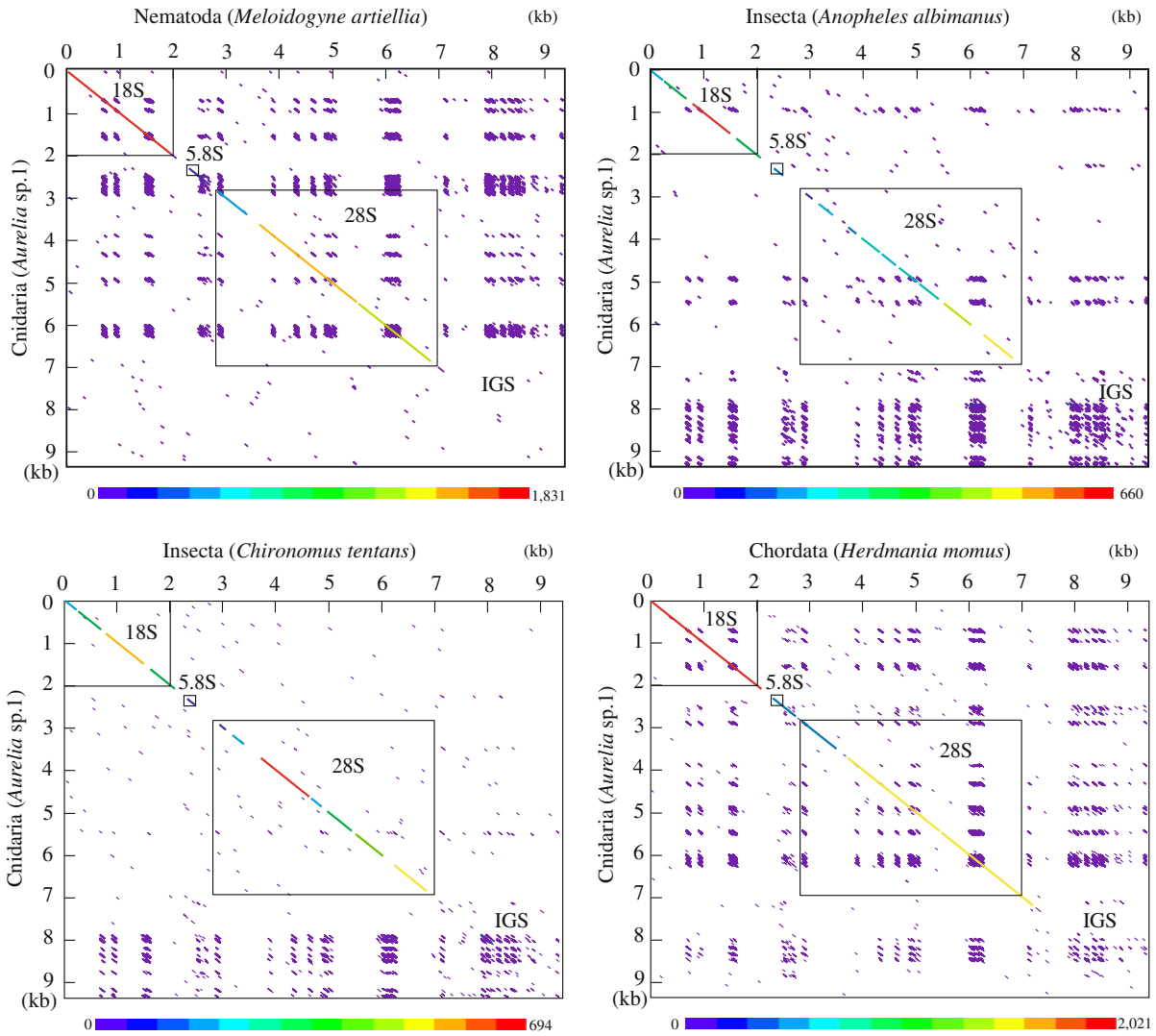

Fig. 4 Dot matrix comparisons of rDNA sequences. Dot plots between *Aurelia* sp.1 with *Meloidogyne artiellia* (Nematoda, AF248477), *Anopheles albimanus* (Arthropoda, L78065), *Chironomus tentans* (Arthropoda, X99212), and *Herdmania momus* (Ascidiacea, X53538) sequences. Color scale bars represent consecutive sequence length of some regions detected similarly between the two sequence pairs. The open boxes in matrixes indicate rDNA coding regions such as 18S, 5.8S, and 28S

(99.7%) was recorded between current *Aurelia* sp. and *Aurelia* sp.1 (AY935214; isolation locality from Miyazu Bay, Japan), followed by 99.4% with *Aurelia* sp.1 (AY935203, California). Interestingly, DNA sequence similarities between our study and those reported by others for *Aurelia* spp. were less than 80% (mostly <70%). This suggests that *Aurelia* sp. studied by us could be assigned to *Aurelia* sp.1. In this context, the current data on *Aurelia* sp.1 were significantly different from those of *A. aurita* (AY935205, 69.2%; AY935206, 71.8%), *A. labiata* (AY935202, 70.8%), and *A. limbata* (AY935215, 79.4%). ITS data from the *Aurelia* sp.1 studied by us were congruent with previous findings (Dawson et al., 2005), where the genus *Aurelia* was distinguished into 13 genotypes that probably represent species.

The tandemly repeated rDNA units of *Aurelia* consisted of coding and non-coding regions. Of them, the IGS of *Aurelia* was more divergent than the other genes (Table 3, Figs. 3, 4) and apparently consisted of three components, i.e. 3′ ETS, NTS, and 5′ ETS.

As a putative RNA polymerase I transcription initiation site, many promoters of the RNA polymerase I are bipartite, consisting of a proximal promoter domain and an upstream control element (UCE) (Chen et al., 2000). In the present study, an internal repeat sequence of "AGG CCG TGT GTG." located at 1,359–1,370 and "AGG CCG AGG GGG." at 1,381–1,392 within the IGS of *Aurelia* was detected (see Figs. 1, 2). These kinds of motifs were consistent with those being proximal promoters and UCE motifs, suggesting a putative proximal domain for RNA polymerase I (Marilley & Pasero, 1996). Taking into consideration both the motif as promoter and transcription probably begins from 20 to 30 bp downstream of the 2nd motif, the 5′ ETS region can be ~190 bp long in *Aurelia* sp.1 Regarding the 5′ ETS, we found a poly(T) tract (TTT ATT TTT ATT) located at positions 21–32 from end side nucleotide of 5′ IGS, which is considered as a signal for termination of the rDNA transcript (Lang et al., 1994; Jeong et al., 1995; Mason et al., 1997). Indeed, a similar pattern of poly(T), which was a tract of "TTT TTT TTT T" and was located at positions 21–32 from end side nucleotide of 5′ IGS, was identified in the IGS of the cnidarian member (*Junceella fragilis*; AF154670). Therefore, the 5′ ETS is likely to be ~30 bp long. The remaining nucleotide sequence belonged to NTS, which was at least 1,350 bp long. Frequently, the NTS region contains many minisatellite sequences, such as "AG", "AC", "AT". However, *Aurelia* NTS has a block of minisatellite, "CTA ACC CTA GCC CTA ACC"-like nucleotide sequence, with five repeats within the IGS. This is the first finding in Cnidaria IGS, representing the number of repeats and nucleotide sequences that may be used as a genetic marker for population history studies.

With regard to IGS length variations, the rDNA IGS of *Aurelia* was short when compared with other eukaryotic IGSs (Table 3). For example, the IGS varies in length from about 2.0 kb (e.g., *Meloidogyne artiellia*, Nematoda, AF248477; *Herdmania momus,* Ascidiacea, X53538), to about 21 kb in mammals and, with very few exceptions, contains significant regions of internal repeating. A pair-wise comparison between *Aurelia* and other eukaryotes showed that low similarities were recorded in all coding and non-coding regions (Table 4). Of them, ITS and IGS showed remarkably low similarities (<47.2%) to the data set. For the comparison of Cnidaria rDNA, we found only one nucleotide sequence of complete rDNA IGS from Anthozoa (*Junceella fragilis* Walker & Bull 1983, AF154670). It contained partial 28S, complete IGS, and partial 18S rDNA. Upon comparison, DNA similarities between them were measured at 80.7% in 285 sites of 28S, at 40.3% in 1,063 sites of IGS, and at 90.3% at 339 sits of 18S. In addition, dot-matrix plots between *Aurelia* sp.1 and four other eukaryotes graphically showed that the sequences of IGS were highly variable with respect to those of other eukaryotic rDNA coding regions (Fig. 4). Furthermore, sequence characteristics such as P–I sites, *p*-distance suggested that the IGS was highly informative of the other rDNA regions (Table 3), although the data included here were sampled from relatively distant relatives such as Nematoda, Arthropoda, and Chordata.

Table 4 Similarity scores (percentage) between *Aurelia* sp.1 and four other eukaryotes

	Aurelia sp.			
	18S	ITS+5.8S	28S	IGS
Anopheles albimanus #L78065	66.5	33.9	56.1	25.4
Chironomus tentans #X99212	65.2	33.9	56.1	29.0
Meloidogyne artiellia #AF248477	69.5	32.9	67.4	29.2
Herdmania momus #X53538	78.6	47.2	70.0	39.8

Analysis of sequence complexity, entropy, and nucleotide distribution provides an efficient way to detect simple sequence repeats in the rDNA. The most simple but frequent case of a low complexity zone is a region of simple sequence repeats. In the current analyses, some signals were detected in the IGS of *Aurelia*. Here, we present plots of complexity distribution along the rDNA using a window size of 100 bases (Fig. 4). By varying the window size, we could obtain a more complete picture of the repeat region. A small-size sliding window reveals relatively short repeats, while a long window may generate long dispersed repeats. In the case of *Aurelia*, a region of low complexity and high entropy corresponded to those of several poly(T) tracts and high AT regions. Overall, analysis of the rDNA revealed that IGS was more informative than the other rDNA regions. These results were highly congruent with the parsimony (Table 3) and dot matrix analyses (Fig. 4).

Conclusion

In this study, we determined for the first time the single unit sequence of the rDNA of the widely distributed moon jelly *Aurelia* rDNA and presented a structural model of nuclear rDNA for molecular comparisons, particularly in jellyfishes. Using molecular analyses, we detected some useful characteristics such as a block of minisatellites, poly(T) tract, and bi-repeat patterns from the *Aurelia* rDNA IGS. Based on these facts, we conclude that rDNA IGS of *Aurelia* sp.1 is highly informative and is potentially useful as a marker for populations to study the global dispersal and expansion of *Aurelia*.

Acknowledgments We would like to thank Drs. Sheikh Raisuddin and Hans-Uwe Dahms for English editing on the manuscript. We are very grateful to Dr. Timothy J. Page and two anonymous reviewers for reading and critical comments. This work was supported by grants of KOSEF NRL (2006) and ETEP (2006) funded to Jae-Seong Lee, and by a grant of KOSEF (2007) funded to Heum Gi Park. Thus work was also supported by a grant from Korea Polar Research Institute (PE08050) funded to Il-Chan Kim.

References

Arai, M. N., 1997. Functional biology of Scyphozoa. Chapman & Hall, New York, NY: 300.

Chen, C. A., D. J. Miller, N. V. Wei, C.-F. Dai & H.-P. Yang, 2000. The ETS/IGS region in a lower animal, the seawhip, *Junceella fragilis* (Cnidaria: Anthozoa: Octocorallia): compactness, low variation and apparent conservation of a pre-rRNA processing signal with fungi. Zoological Studies 39: 138–143.

Collins, A. G., P. Schuchert, A. C. Marques, T. Jankowski, M. Medina & B. Schierwater, 2006. Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models. Systematic Biology 55: 97–115.

Dawson, M. N., 2003. Macro-morphological variation among cryptic species of the moon jellyfish, *Aurelia* (Cnidaria: Scyphozoa). Marine Biology 143: 369–379.

Dawson, M. N. & D. K. Jacobs, 2001. Molecular evidence for cryptic species of *Aurelia aurita* (Cnidaria, Scyphozoa). Biological Bulletin 200: 92–96.

Dawson, M. N. & L. E. Martin, 2001. Geographic variation and ecological adaptation in *Aurelia* (Scyphozoa: Semaeostomeae): some implications from molecular phylogenetics. Hydrobiologia 451: 259–273.

Dawson, M. N., A. Sen Gupta & M. H. England, 2005. Coupled biophysical global ocean model and molecular genetic analyses identify multiple introductions of cryptogenic species. Proceedings of the National Academy of Sciences of the United States of America 102: 11968–11973.

Drouin, G., J. D. Hofman & W. F. Doolittle, 1987. Unusual ribosomal RNA gene organization in copepods of the genus *Calanus*. Journal of Molecular Biology 196: 943–946.

Hassouna, N., B. Michot & J.-P. Bachellerie, 1984. The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. Nucleic Acids Research 12: 3563–3583.

Hillis, D. M. & M. T. Dixon, 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. Quarterly Review of Biology 66: 411–453.

Jeong, S. W., W. H. Lang & R. H. Reeder, 1995. The release element of the yeast polymerase I transcription terminator can function independently of Reb1p. Molecular and Cellular Biology 15: 5929–5936.

Ki, J.-S. & M.-S. Han, 2007. Informative characteristics of 12 divergent domains in complete large subunit rDNA sequences from the harmful dinoflagellate genus, *Alexandrium* (Dinophyceae). Journal of Eukaryotic Microbiology 54: 210–219.

Ki, J.-S., D.-S. Hwang, K. Shin, W. D. Yoon, D. Lim, Y. S. Kang, Y. Lee & J. -S. Lee, 2008. Recent moon jelly (*Aurelia* sp.1) blooms in Korean coastal waters suggest global expansion: examples inferred from mitochondrial COI and nuclear ITS–5.8S rDNA sequences. ICES Journal of Marine Science 65: 443–452.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111–120.

Kramp, P. L., 1961. Synopsis of the medusae of the world. Journal of the Marine Biological Association of the United Kingdom 40: 337–342.

Kumar, S., K. Tamura, I. B. Jakobsen & M. Nei, 2001. MEGA3: Molecular Evolutionary Genetics Analysis software. Bioinformatics 17: 1244–1245.

Lang, W. H., B. E. Morrow, Q. Ju, J. R. Warner & R. H. Reeder, 1994. A model for transcription termination by RNA polymerase I. Cell 79: 527–534.

Lee, J.-S., 2000. The internally self-fertilizing hermaphroditic teleost *Rivulus marmoratus* (Cyprinodontiformes, Rivulidae) β-actin gene: amplification and sequence analysis with conserved primers. Marine Biotechnology 2: 161–166.

Marilley, M. & P. Pasero, 1996. Common DNA structural features exhibited by eukaryotic ribosomal gene promoters. Nucleic Acids Research 24: 2204–2211.

Mason, S., M. Wallisch & I. Grummt, 1997. RNA polymerase I transcription termination: similar mechanisms are employed by yeast and mammals. Journal of Molecular Biology 268: 229–234.

Mathews, D. H., J. Sabina, M. Zuker & D. H. Turner, 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. Journal of Molecular Biology 288: 911–940.

Mayer, A. G., 1910. Medusae of the World, Vol. 3. The Scyphomedusae. Carnegie Institution of Washington, Washington, DC: 603–630.

Schlötterer, C., 1998. Ribosomal DNA probes and primers. In Karp, A., P. G. Isaac & D. S. Ingram (eds), Molecular

Tools for Screening Biodiversity. Chapman & Hall, London: 267–276.

Schroth, W., G. Jarms, B. Streit & B. Schierwater, 2002. Speciation and phylogeography in the cosmopolitan marine moon jelly, *Aurelia* sp. BMC Evolutionary Biology 2: 1–10.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31: 3406–3415.