J Mol Evol (1988) 27:336–340

Journal of Molecular Evolution

Studies on the Sites Expressing Evolutionary Changes in the Structure of Eukaryotic 5S Ribosomal RNA

Guo-rong Qi, Gong-jie Cao, Peng Jiang, Xiao-li Feng, and Xiang-rong Gu

Shanghai Institute of Biochemistry, Academia Sinica, 320 Yue-Yang Road, Shanghai 200031, China

Summary. We have determined the complete sequences of 5S rRNAs from a lamprey (*Lampetra reissneri*), a lancelet (*Branchiostoma belcheri*), silkworms (*Philosamia cynthia ricini, Bombyx mori, Antheraea pernyi*), and a silkworm hybrid (artificially fertilized hybrid species of *Philosamia cynthia ricini* ♂ × *Bombyx mori* ♀), as well as those of cotton seeds (*Gossypium hirsutum* L.). Having compared more than 170 eukaryotic 5S rRNAs of which seven sequences have been determined by our group as mentioned above, we have found that the "evolutionary sites" that exist at special locations in these structures are closely related to the evolution of eukaryotes. The changes proceed step by step in an orderly way, i.e., the change in nucleotide residues of the "evolutionary sites" depends on the order of the evolution of the species and shows group-specific patterns.

Key words: 5S rRNA — Nucleotide sequence — Secondary structure — Evolutionary sites

Introduction

Protein molecules such as hemoglobin and cytochrome *c* have, for a long time, been used to study evolution or phylogenesis at the molecular level. Now, nucleic acid molecules, particularly the small RNAs, are targeted subjects because they exist in every organism and are the direct products of gene expression. 5S rRNA is a prime subject because its entire sequence is easy to determine. Osawa and Hori deduced an empirical formula from the difference in the number of nucleotide residues between 5S rRNAs of different species, and used this to calculate the order of phylogenetic branching and relative evolutionary distance. Consequently, a model for a molecular phylogenetic tree was proposed (Osawa and Hori 1979). However, the above-mentioned molecular evolution studies compared only the random differentiation of residues of amino acids in proteins or nucleotides in nucleic acids. Sometimes 5S rRNAs of the same origin have different sequences. Take for example: two nucleotide differences between two *Bombyx mori* 5S rRNAs (Komiya et al. 1981; this paper); four differences between two *Philosamia cynthia ricini* 5S rRNAs (Gu et al. 1982; Cao et al. 1983); 82% sequence similarity in two 5S rRNAs from *Emplectonema gracile* (Kumazaki et al. 1983); or the common occurrence of heterogeneity in the 5S rRNAs between oocytes and somatic cells (Delihas and Andersen 1982). We must therefore not consider that these rRNAs belong to different evolutionary stages.

Since 1979 we have determined the sequences of 5S rRNAs from seven eukaryotic organisms: a lamprey [*Lampetra reissneri* (Cyclostomata; one of the lowest of vertebrates)], a lancelet [*Branchiostoma belcheri* (Cephalochordata; one of the highest of invertebrates)], *Gossypium hirsutum* L., and four different species of silkworm. Having compared more than 170 eukaryotic 5S rRNAs, including the seven sequences mentioned above, we have found that the evolutionary sites as we call them, which exist at special sites in the structure of 5S rRNAs (signature analysis; Delihas and Andersen 1982), are closely related to the evolution of eukaryotes.

Offprint requests to: G-r Qi

GCCUACGACCAUAUCACCCUGAAUACGCCCGAUCUCGUCCGAUCUCGGAAGCUAAGCAGGGUCGAGCCUGGUUAGUACUU

GCCUACGGCCAUAUCACGUUGAGUACACCCGAUCUCGUUCGAUCUCGGAAGUUAAGCAACGUCGAGUCCGGUCAGUACUU

GCCAACGUCCAUACCACGUUGAAAACACCGGUUCUCGUCCGAUCACCGAAGUUAAGCAACGUCGGGCGCGGUCAGUACUU

GACAACGUCCAUACCACGUUGAAAACACCGGUUCUCGUCCGAUCACCGAAGUUAAGCAACGUCGGGCGCGGUCAGUACUU

GCCAACGUCCAUACCAUGUUGAAUACACCGGUUCUCGUCCGAUCACCGAAGUCAAGCAACAUCGGGCGUAGUCAGUACUU

GCCAACGUCCAUACCAUGUUGAAUACACCGGUUCUCGUCCGAUCACCGAAGUCAAGCAACAUCGGGCGUAGUCAGUACUU

GGUUCAAUCAUACCGACACUAAUGCACCGGAUCCCAUCAGAACUCCGCAGUUAAGCGUGCUUGGGCGAGAGCAGUACUAG

GGAUGGGAGACCGCCUGGGAAUACCAGGUGUUGUAGGCUU --- Lampetra reissneri (Jiang et al. 1986)*

GGAUGGGAGACCGCCUGGGAACACCGGGAGUUGUAGGCAU --- Branchiostoma belcheri

GGAUGGGUGACCGCCUGGGAACACCGCGUGUUGUUGGCUU --- Philosamia cynthia ricini (Cao et al. 1983)

GGAUGGGUGACCGCCUGGGAACACCGCGUGACGUUGGCUU --- Antheraea pernyi

GGAUGGGUGACCGCCUGGGAACACUACGUGAUGUUGGCUU --- Bombyx mori (757)

GGAUGGGUGACCGCCUGGGAACACUACGUGAUGUUGGCUU --- Silkworm hybrid (Philosamia ♂ X Bombyx ♀)

GAUGGGUGACCUCCUGGGAAGUCCUCGUGUUGAACCCU --- Gossypium hirsutum L. (Qi et al. 1985)

Fig. 1. The primary structure of seven 5S rRNAs. *The *Lamptera* sequence that we published in 1986 is identical to that of another lamprey species, *Entosphenus japonicus*, as determined by other laboratories (Komiya et al. 1986).

Materials and Methods

Materials. Silkworms (*Antheraea pernyi, Bombyx mori*) and a silkworm hybrid (artificially fertilized species from hybridization of *Philosamia cynthia ricini* ♂ × *B. mori* ♀) were cultured by the Kuangxi Institute of Sericulture, Nanning. A lancelet (*Branchiostoma belcheri*) was purchased from Xia-men and a lamprey (*Lampetra reissneri*) was a gift from Dr. Chunsheng Guo of Harbin Normal University, Harbin. For chemical reagents and enzymes used in sequence determination, refer to previous papers (Cao et al. 1983; Qi et al. 1985).

Methods. The nucleotide sequences of 5S rRNA were determined by means of partial chemical or/and enzymatic degradation techniques (Peattie 1979; Donis-Keller 1980) with modifications described by Cao et al. (1983) and Qi et al. (1985). After determining the primary structure of 5S rRNA, we constructed the secondary structure according to the model proposed by De Wachter et al. (1982). By comparing the secondary structures of all known eukaryotic 5S rRNAs, we looked for regularity in the relationship between nucleotide change at special sites in the structure and correlated this with the classical taxonomic positioning of the organisms (signature analysis; Delihas and Andersen 1982).

The method of drawing secondary structures advocated by Erdmann et al. (1985) was used. In Fig. 2 the conserved sites possessed by all eukaryotes are represented by common nucleotide letters A, C, G, U, and the semiconserved sites by special letters such as R, Y, or S (for details see notes to Fig. 2). The substitution sites (with more than three nucleotide changes) are shown by open circles, and the sites expressing evolutionary changes in structure are marked with star-circles and Roman numerals (see Fig. 2).

Results and Discussion

The Primary Structures of 5S rRNAs

The primary structures of seven 5S rRNAs are listed in Fig. 1. The structures of 5S rRNAs from *Philosamia cynthia ricini, Gossypium hirsutum* L., and *Lampetra reissneri* have already been published. The sequences from *Branchiostoma belcheri, Antheraea pernyi, Bombyx mori* 757, and a silkworm hybrid have been completed.

The Relationship between Evolution and Structure of 5S rRNAs in Eukaryotes

Figure 2 illustrates the similarity and diversity of 5S rRNA structure in eukaryotes, following the pattern proposed by De Wachter et al. (1982). It appears that the evolutionary sites are mainly concentrated on stem C, and also on stem B. Some examples of the base change in different species are as follows: evolutionary site I (residue No. 9 in Fig. 2)—U for higher plants, C for animals, Eumycota and Myxomycota; sites IV_1–IV_2 (Nos. 27–52)—G–C for vertebrates and A–U for other eukaryotes; site VIII (No. 35)—U for Metazoa and Mesozoa, with C for other eukaryotes, etc. Similar phenomena demonstrating changes in other evolutionary sites such as

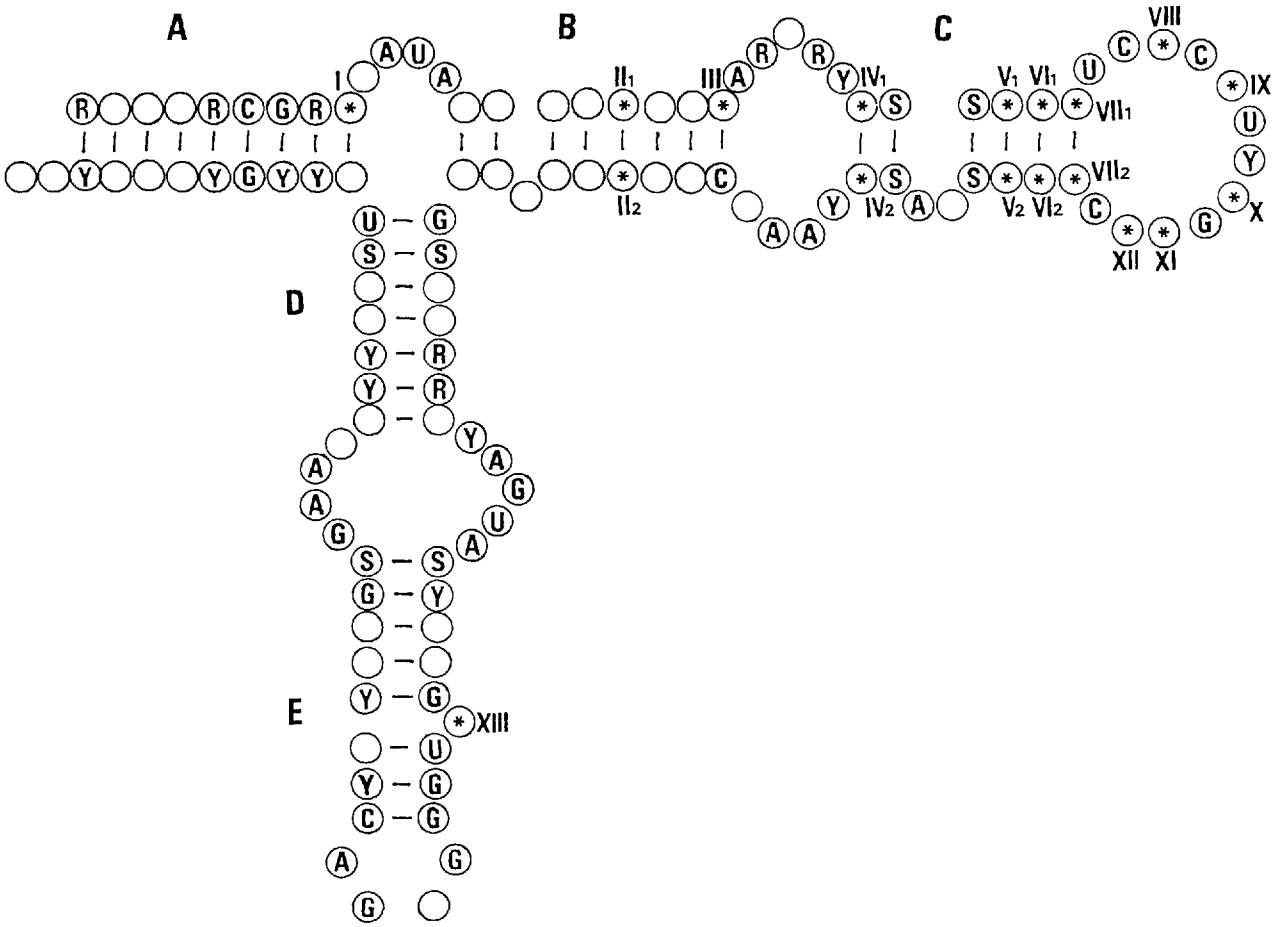

Fig. 2. Similarity and diversity of 5S rRNA structures in eukaryotes. *Notes:* (1) A, B, C, D, and E out of the structure represent the stems of 5S rRNA structure; (2) A, C, G, and U in the circle—the common nucleotides; (3) R, Y, and S in the circle—purine (A or G), pyrimidine (C or U), and G or C nucleotides separately; (4) open circles are the substitution sites in which there are more than three nucleotide changes; (5) the star-circles represent evolutionary sites and are marked with Roman numerals.

Table 1. The evolutionary sites of 5S rRNA in animals

Classical taxonomy	5S rRNA number	Evolutionary sites (location of residues)											
		I (9)	XI (42)	III (21)	VI_1–VI_2 (31–46)	VIII (35)	XIII (83)	XII (43)	IX (37)	V_1–V_2 (30–47)	II_1–II_2 (18–60)	IV_1–IV_2 (27–52)	X (40)
Protozoa	11	C	C A	U, A G	C–G G–C	C	G	A U	A G	A–U G–C	G–C	A–U	C
Mesozoa	1	C	A	G	G–C	U	C	C	A	C–G	G–C	A–U	C
Metazoa													
Other invertebrates	48	C	A	G	G–C	U	A	U	G	C–G A–U G–C	G–C	A–U	C
Hemichordata (*Saccoglossus kowalevskii*)	1	C	A	G	G–C	U	A	U	G	G–C	G–C	A–U	C
Urochordata (*Halocynthia roretzi*)	1	C	A	G	G–C	U	A	U	G	C–G	A–U	A–U	C
Cephalochordata (*Branchiostoma belcheri*)	1	C	A	G	G–C	U	A	U	G	C–G	G–C	A–U	C
Vertebrates													
Cyclostomata (*Lampetra reissneri*)	3	C	A	G	G–C	U	A	U	G	C–G	C–G	G–C	C
Pisces	5	C	A	G	G–C	U	A	U	G	C–G	C–G	G–C	C U
Amphibia, Reptilia, Aves, Mammalia	9	C	A	G	G–C	U	A	U	G	C–G	C–G	G–C	U

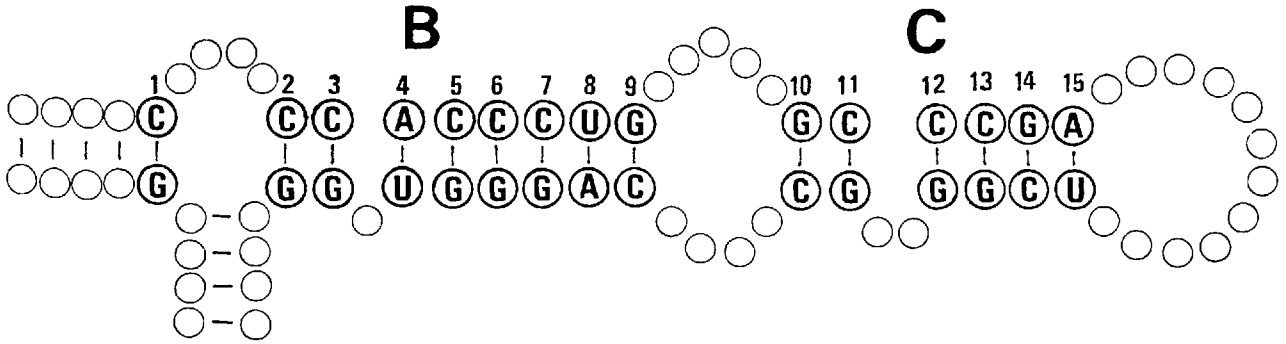

Fig. 3. The structural features of stems B and C of vertebrate 5S rRNA. *Notes:* See also notes 1 and 2 in Fig. 2. Site 1—U–G pairing in plants; sites 2 and 3—two U–G pairings not existing in all organisms; sites 4, 8, and 15—three A–U pairings in nearly all organisms including vertebrates; sites 5, 6, and 7—three C–G pairings take up very few C–G continuous pairings (about 10%) except in vertebrates; site 9—no pairing in plants; and site 13—G–C pairing except in vertebrates and protochordates.

residue Nos. 21, 37, 40, 42, 43, and 83, as well as bp Nos. 16–80, 30–74, 31–46, and 32–45, etc., have also been observed.

There seems to be no regularity at first glance of Fig. 2, but the regularity becomes clear when we arrange the evolutionary sites in accordance particularly with the classical taxonomic positions of animals from lower to higher species. Table 1 shows the results indicating changes of evolutionary sites; some occur at early periods of evolution of animals and others occur at later periods. This means that the base change in the evolutionary sites is on the order of the evolution of species. Though we do not know the actual significance of the order of change in evolutionary sites, it is interesting to point out that it is basically consistent with the results of classical taxonomic positioning of animals. Clearly, each of the following organisms occupies a specific position in Table 1—*Dicyma misakiense* (Mesozoa), *Saccoglossus komalevskii* (Hemichordata), *Halocynthia roretzi* (Urochordata), *Branchiostoma belcheri* (Cephalochordata), and *Lampetra reissneri* (Cyclostomata).

There are a few exceptions in Table 1 and they are always found in lower organisms. However, these exceptions do not seem to influence the conclusions listed in Table 1. Although the true biological significance of the above-described orderly changes of nucleotide residues at evolutionary sites is still unknown, the results may be used as a probe to identify an organism and its location in evolution by means of sequence analysis followed by reference to its regularity of change at evolutionary sites. For example, some taxonomists do not consider *Saccoglossus komalevskii* to be a protochordate, which is in agreement with the results shown in Table 1 because the evolutionary site V_1–V_2 of *Saccoglossus komalevskii* is a G–C pair (not a C–G pair), similar to invertebrates but distinct from the protochordates.

Part of the 5S rRNA structures of vertebrates that are derived from Fig. 2 and Table 1 are listed in Fig. 3. In addition to the three A–U pairs possessed by all eukaryotes, there are 11 G–C pairs from the total 14 base pairings in stems B and C of 5S rRNA in vertebrates, which is much higher than in other species. What role did the high GC content in this area play on the 5S rRNA secondary structure of vertebrates? This deserves further investigation.

We have also tried to trace the evolutionary sites of 5S rRNAs in plants. We have not found much sequence similarity in evolutionary sites after comparing the structures of the 5S rRNAs from plants and Eumycota. Therefore it seems reasonable for some taxonomists to take the Eumycota out of plants and reclassify them as two special kinds of organisms.

Acknowledgment. We thank Prof. Wang TP of our Institute for critical reading of the manuscript.

References

Cao GJ, Li WQ, Gu XR, Zhu LF, Qi GR, Wang DB (1983) The complete sequence of 5S rRNA from the posterior silkgland of the silkworm *Philosamia cynthia ricini*. Acta Biochim Biophys Sin 15:405–413

Delihas N, Andersen J (1982) Generalized structures of the 5S ribosomal RNAs. Nucleic Acids Res 10:7323–7344

De Wachter R, Chen MW, Vandenberghe A (1982) Conservation of secondary structure in 5S rRNAs: a uniform model for eukaryotic, eubacterial, archaebacterial and organelle sequences is energetically favourable. Biochimie 64:311–329

Donis-Keller H (1980) Phy M: an RNase activity specific for U and A residues in RNA sequence analysis. Nucleic Acids Res 8:3133–3142

Erdmann VA, Wolters J, Huysmans E, De Wachter R (1985) Collection of published 5S, 5.8S and 4.5S ribosomal RNA sequences. Nucleic Acids Res 13:r105–r141

Gu XR, Nicohgosian K, Cedergren RJ (1982) 5S rRNA sequence from the *Philosamia* silkworm: evidence for variable

evolutionary rates in insect 5S rRNA. Nucleic Acids Res 10: 5711–5716

Jiang P, Qi GR, Guo CS, Wang QS (1986) Structure analysis of ribosomal 5S rRNA in *Lampetra reissneri*. Acta Biochim Biophys Sin 18:391–395

Komiya H, Kawakami M, Takemura S (1981) Nucleotide sequence of 5S rRNA from the posterior silkglands of *Bombyx mori*. J Biochem 89:717–722

Komiya H, Hasegawa M, Takemura S (1986) Differentiation of oocyte- and somatic type 5S rRNAs in animals. J Biochem 100:369–374

Kumazaki T, Hori H, Osawa S (1983) The nucleotide sequences of 5S rRNA from two ribbon worms: *Emplectonema gracile* contains two 5S rRNA species differing considerably in their sequences. Nucleic Acids Res 11:7141–7144

Osawa S, Hori H (1979) Molecular evolution of ribosomal components. In: Chambliss G et al. (eds) Ribosomes—structure, function and genetics. University Park Press, Baltimore, p 333

Peattie DA (1979) Direct chemical method for sequencing RNA. Proc Natl Acad Sci USA 76:1760–1764

Qi GR, Cao GJ, Zhu LF, Feng XL, Zhang YP, Zheng ZR (1985) The complete sequence of the 5S rRNA from cotton seeds. Acta Scientiarum Universitatis Sunyatseni, No. 2, pp 65–72

Received January 19, 1987/Revised January 27, 1988