Genomic organization of the rDNA cistron of the teleost fish *Cyprinus carpio*

MARÍA INÉS VERA, ALFREDO MOLINA, RODRIGO PINTO, MAURICIO REYES, MARCO ÁLVAREZ, ERWIN KRAUSKOPF, CLAUDIA QUEZADA, JORGE TORRES and MANUEL KRAUSKOPF

Millennium Institute for Fundamental and Applied Biology, Universidad Andrés Bello, Santiago, CHILE

ABSTRACT

The seasonal adaptation of the teleost *Cyprinus carpio* to the cyclical changes of its habitat demands physiological compensatory responses. The process involves profound nucleolar adjustments and remarkable changes in rRNA synthesis, which affects ribosomal biosynthesis. In this context, we have demonstrated that the synthesis of several proteins involved in ribosomal biogenesis as protein kinase CK2, ribosomal protein L41 and nucleolin, as well as U3 snoRNP, are differentially regulated in summer-acclimatized carp compared to the cold-season adapted fish.

To understand the mechanisms involved in the seasonal regulation of rRNA gene transcription, we have been studying the carp rDNA cistron structure. Because the cis-elements that regulate the expression of the tandem organized ribosomal genes are located in the non-transcribed intergenic spacer (IGS), we analyzed the primary structure of the carp rDNA gene IGS. The gene organization is similar to that described from other vertebrate species, including numerous repetitive sequences, the transcription start site, and some potential cis-elements such as ribosomal enhancers, proximal terminator and transcriptional terminators.

Ribosomal DNA is a remarkable case of gene duplication and has been used as a model to test the concerted evolution theory. We performed sequence comparison analyses of 18S rRNA coding sequences from carp with different species, data with which an unrooted phylogram was constructed.

Key terms: ribosomal gene, rRNA; rDNA; fish; carp

INTRODUCTION

The process of protein synthesis, which is mediated by ribosomes, is essential for cell growth, proliferation, and adaptation to changing environments. The ribosomal biogenesis is a very complex process that involves the synthesis and assembly of four different ribosomal RNA (rRNA) molecules and of 80 different proteins that takes place in the nucleolus of eukaryotic cells. The transcription of a 45S rRNA precursor (pre-rRNA) takes place in this nuclear compartment, along with its processing and maturation in the 18S, 5.8S and 28S rRNAs and the assembly of the pre-ribosomal particle. The ribosome synthesis gives rise to the characteristic ultra-structural organization of the nucleolus into domains that contain a minimum of three distinctive components: the fibrillar center, the dense fibrillar component, and the granular component (Lazdins et al., 1997). Nucleolar structure is largely dependent upon the process of ribosome assembly. The nucleolus is disassembled during mitosis, at which time pre-rRNA transcription and processing are suppressed (Hernandez-Verdun et al., 2002). Actinomycin D treatment affects the structure and integrity of the nucleolus in HeLa cells, resulting in a segregation of nucleolar components and rRNA transcription repression (Schofer et al., 1996), along with changes in the relative abundance of a subset of nucleolar proteins (Andersen et al., 2002). It is
known that in the human meiotic oocyte, the nucleolus shows spontaneous segregation of its components due to temporary inactivation of the ribosomal genes expression (Mirre et al., 1980).

The habitat of the carp (Cyprinus carpio) undergoes seasonal changes that compel the fish to develop adaptive strategies for sensing the cyclical fluctuation of the physical parameters (e.g. temperature, photoperiod) and to transduce them into molecular signals through mechanisms that involve the neuroendocrine system (Figueroa et al., 1994; Kausel et al., 1998; 1999; López et al., 2001). One of the most outstanding cellular features of this adaptive process is the reorganization of the nucleolus architecture. In the cells of winter-adapted fish, the nucleolar structure is segregated and the fibrillar components are densely packed and separated from the granular component and the associated chromatin (Sáez et al., 1984; Vera et al., 1993). Concurrently, the transcription of ribosomal genes by RNA polymerase I and the processing of the rRNA precursors decrease dramatically during the cold season (Vera et al., 1993). During summer, however, the nucleolar structure reflects the active synthesis and processing of rRNA, e.g. the fibrillar and granular components are intermingled (Sáez et al., 1984; Vera et al., 1993).

The natural segregation of the carp nucleolar components and the down regulation of rRNA transcription during the adaptation of the carp to the cold season are analogous to the blocking of rRNA synthesis and nucleolar segregation obtained in vitro with Actinomycin D. Thus, the carp acclimatization constitutes a valuable model to study the regulation of ribosome biogenesis, which is closely synchronized at multiple steps, including the regulation of pre-rRNA synthesis at different stages of the process, e.g. chromatin remodeling, transcriptional activation, initiation, elongation and termination (Leary and Huang, 2001). All these regulatory steps involve interactions between cis-acting elements of rDNA genes and trans-acting protein factors. To approach the study of the regulation of rRNA synthesis during the seasonal adaptation of the carp, it is necessary to know the structure of the gene that encodes for the rRNA precursor and the cis elements located at the regulatory region of the carp ribosomal cistron. In this report we describe the primary sequence arrangements of the carp ribosomal cistron potentially involved in the initiation and termination of the transcription, and complement the previous carp rDNA sequences we have described previously (Vera et al., 1997; Joseph et al., 1999).

**MATERIALS AND METHODS**

**Animal and tissues:**

Male carp were captured and maintained at summer- (20-22°C) and winter- temperatures (8-10°C), and the tissues dissected and frozen as described (Vera et al., 1993).

**Genomic library screening**

A 72bp fragment, coding for a partial sequence of the 5.8S rRNA (Vera et al., 1997) was used to screen a 1FIX II carp genomic library (Stratagene, USA), which yielded 100 hybridizable clones under stringent conditions (Grossberger, 1987). Six recombinant 1-clones were chosen to be purified by three additional screenings. Two (1CRR11 and 1CRR13 with inserts of 14kb and 16kb, respectively) were analyzed by restriction mapping, and their nature confirmed by Southern blot using the 5.8S probe. In addition, a 177bp fragment corresponding to the D3 region of the carp 28S rRNA (Michot et al., 1990) was used as probe in Southern blot analyses.

Figure 1 depicts the sub-cloning strategies used for sequence analyses of both 1CRR11 and 1CRR13 clones. A 5kb Xba I fragment from the 1CRR11 clone, hybridizing with the 5.8S and 28S probes, was sub-cloned (pCRR5X) into the pUC19 vector and sequenced. A 2kb Eco RI-Pst I fragment from the 1CRR13 clone, overlapping the 5’ end of the pCRR5X clone
was sub-cloned (p13CRR2ep) into the pUC19 vector and sequenced. A 6kb Hind III-Eco RI fragment from the 1CRR13 clone, flanking the 5’ end of the p13CRR2ep, was subcloned (p13CRR6.6he) into pUC19 vector and partially sequenced. A 1.5kb Eco RI-Hind III fragment, flanking the 5’ end of the p13CRR6.6he, was sub-cloned (p13CRR1.5eh) into pUC19 vector and sequenced. Finally, a 1kb Sal I-Eco RI fragment, flanking the 5’ end of the p13CRR1.5eh, was sub-cloned (p13CRR1se) into pUC19 vector and sequenced. The p13CRR2ep clone was sub-cloned (p13CRR0.7sn) into the pBS SK+ vector. From the p13CRR0.7sn clone linearized with Pst I, using T3 RNA polymerase (Promega, USA), a riboprobe was generated by in vitro transcription and simultaneously radio-labeled with [α-32]CTP. Total kidney RNA (2µg) extracted from both winter- and summer–acclimatized carp (Sambrook and Russell, 2001) were hybridized with the riboprobe (100.000cpm) and treated with a nuclease mix (RNAse A/RNAse T1). The RPA products were fractionated in a denaturing 7% polyacrylamide gel and exposed to an X-Omat film (Kodak, USA).

Transcription start site

To determine the transcription initiation site, we performed a ribonuclease protection assay (RPA) using the RPA III kit (Ambion, USA) according to the manufacturer’s instructions. Briefly, a 743bp Sma I-Not I fragment from the p13CRR2ep clone was sub-cloned (p13CRR0.7sn) into the pBS SK+ vector. From the p13CRR0.7sn clone linearized with Pst I, using T3 RNA polymerase (Promega, USA), a riboprobe was generated by in vitro transcription and simultaneously radio-labeled with [α-32]CTP. Total kidney RNA (2µg) extracted from both winter- and summer–acclimatized carp (Sambrook and Russell, 2001) were hybridized with the riboprobe (100.000cpm) and treated with a nuclease mix (RNAse A/RNAse T1). The RPA products were fractionated in a denaturing 7% polyacrylamide gel and exposed to an X-Omat film (Kodak, USA).

Phylogenetic analysis

The 18S rRNA genes from table II were aligned using ClustalX (Thompson et al., 1997) with default gap penalties. The alignment was subsequently optimized.
manually. To build the phylogenetic tree, an analysis based on the Neighbor-Joining method was performed using the software MEGA v2.1 (Kumar et al., 2001) with 5,000 replicates using only bootstrap values over 50.

RESULTS AND DISCUSSION

The down regulation in the synthesis and processing of rRNA (Vera et al., 1993) and in the transcription of the ribosomal protein L41 during the adaptation of the carp to the cold season environment conditions (Molina et al., 2002) suggest the existence of a complex and coordinated mechanism to modulate the ribosome biogenesis during the acclimatization process. In this study, we described the carp rRNA cistron sequence in order to dissect the structure and organization of its most important features that are potentially involved in the carp seasonal regulated rRNA expression.

The rRNA genes are represented in multiple copies throughout the genome of eukaryotic cells, varying in size and copy numbers from species to species (Long and Dawid, 1980; Cortadas and Pavon, 1982). These genes are clustered among many chromosomal loci, ordered in tandem in a “head to tail” manner. Each copy encodes for a polycistronic RNA precursor which is processed into three mature rRNAs (18S, 5.8S, and 28S), separated by two internal transcribed spacers (ITS-1 and ITS-2), and flanked by two external transcribed spacers (5′-ETS and 3′-ETS). Between each coding unit there is a non-transcribed DNA segment called an intergenic spacer (IGS) and containing a number of repeated sequences, some of which enhance the rDNA promoter activity (Paule, 1998).

In most of eukaryotic species, the length diversity of rDNA genes is represented by the IGS variation, ranging from 2kb in yeast to over 20kb in mammals (Paule, 1998). Our results show that the complete carp rRNA gene is contained between both genomic clones (1CRR11 and 1CRR13). The size of the carp rDNA cistron is approximately 15kb, of which 8.3kb correspond to the non-coding sequence (IGS) and 6.7kb to the transcribed region (GenBank accession numbers AF133089 and AY260899).

Among different species, the nucleotide sequence of IGS represents the most variable region of the rDNA genes and contains the cis-elements that regulate transcription mediated by RNA polymerase I (Fath et al., 2001). The regulatory sequences of rDNA promoters are arranged in the same order and general position relative to the +1 transcription site, even though they are quite divergent in sequence (Paalman et al., 1995). RPA analyses of the carp rDNA gene localize the transcription start site at an A residue 418bp upstream of the 18S rRNA coding sequence (Figs. 2 and 3). Secondary protected fragments were found at +135, +141 and a strong-labeled signal at +165 that could correspond to partially processed transcripts (Fig. 2). The protected fragments observed from RPA analyses performed with RNA extracted from summer- and winter-adapted carp presented the same size. Nevertheless, stronger signals were found in summer-acclimatized carp RPA products, suggesting that seasonal changes in rRNA transcription are not mediated by a potential alternative promoter and/or by a differential processing of the precursor at the 5′ETS. These results are consistent with the diminished transcription and processing of rRNA previously reported by in vitro transcription experiments in pituitary glands from winter-adapted carp (Vera et al., 1993).

Despite the lack of conservation in the rDNA regulatory elements, using in vitro assays to study transcriptional activity in different species reveals a core promoter within approximately –40 to +5 with respect to the transcription start site, which is sufficient to drive transcription initiation (Doelling and Pikaard, 1995; Radebaugh et al., 1997). In mammals and Xenopus rDNA genes, a second regulatory element is the upstream promoter element (UPE), located around 150bp upstream from the core promoter (Paule, 1998). The upstream binding factor (UBF) binds UPE and enhances the rRNA transcription by recruiting the core promoter factors (Xie and Rothblum, 1992). An alignment of the
Figure 2
Transcription initiation site determined by ribonuclease protection assay (RPA). Lane 1 shows the riboprobe; Lane 2 shows negative control (without RNA); Lane 3 shows the RPA product of RNA from summer-acclimatized carp; and Lane 4 the RPA product of RNA from winter-acclimatized carp. A non related sequence used as a molecular size marker is on lines denoted by the number 5. Sequence comparison of the region surrounding the transcription start site of different species is shown at the bottom.
Figure 3
Schematic representation and nucleotide sequence of the region upstream of the 18S rRNA including the 5'ETS and a partial sequence of the IGS. (A) The localization of the repeat sequences represented by boxes, transcription start site (arrow) and 18S rRNA are schematized on the top. (B) The sequence numeration is related to the transcription start site denoted by italic bold uppercase letter. The 5'R115 and the 5'R330 repeats are shadowed; the region comprising the 5'R120 repeats is underlined and the repeats alternate between bold and normal letters. Dark shadowed letters denote putatives T₀ and T₀'.

region upstream to the carp +1 rDNA with other species was unable to detect major transcriptional controlling cis-elements, e.g., the core promoter or the upstream promoter (UPE). In most eukaryotic species, the region surrounding the transcription start site is the only conserved sequence (Perna et al., 1992). Surprisingly, the alignment with X. laevis and A. thaliana depicts a core sequence identity up to 6bp downstream of the transcription start site (Fig. 2). The remaining IGS contains repeated sequences that regulate the transcription developed through evolution by repetitive duplication and truncation (Paule and White, 2000 and references within).

The rDNA promoter sequence repeats, called ribosomal enhancers, contain cis-elements that stimulate transcription mediated by the core promoter. The size and number of those sequence repeats present a high variability among different species, 60 and 81bp repeats in Xenopus (Reeder, 1984), 240 and 330bp repeats in Drosophila (Grimaldi and Di Nocera, 1988), 140bp repeats in mouse (Kuhn et al., 1990), several different sized repeats in rat (Cassidy et al., 1986), a 140bp repeat in Acanthamoeba (Paule et al., 1991), and even a single 200bp in yeast (Elion and Warner, 1986). In some species, notably in Xenopus, this enhancer activity passes through the binding of transcription factors which stimulate the transcriptional machinery in the core promoter (Caudy and Pikaard, 2002). Additionally, it has been proposed that if those repeats are too distant from the +1 site, they inhibit transcription because they compete with the promoter for transcription factors as UBF (Pape et al., 1989; Pikaard et al., 1989).

In the carp IGS we found several different repeat elements, with notorious variations in sequence and size (Figs. 3 and 4). Two direct 330bp repeats at positions -443 and -1064 (5'R3301 and 5'R3302, respectively); nine 120 bp direct repeats were detected at positions -2177, -2064, -1937, -1811, -1686, -1563, -1423, -1310 and -1202 (5'R1201 to 5'R1209); and two direct 115bp repeats at positions -3733 and -2694 (5'R1151 and 5'R1152, respectively). Downstream of the 3’end of carp 28S rRNA we found five repetitions of the sequence GGGTGACCAGTGG, potentially representing the transcription terminators (3’T1 to 3’T5), (Figure 4). Four of those repeats are included in two larger sequence repeats (179bp) located at 199bp.
and 580 bp from the 3'-end of the 28S (3'R_{1801} and 5'R_{1802}). The same sequence repeat was found upstream of the +1 embodied in the direct repeats 5'R_{3301} and 5'R_{3302}, probably corresponds to the proximal terminators T₀ and T₀' (Figure 3). Another sequence repeat of 146 bp was found at 1094, 1141, and 1387 bp downstream from the 28S rRNA (3'R_{1461} to 3'R_{1463}). Half of this repeat was also found 1044 bp downstream of the 28S, (h3'R_{1464}) at the beginning of the repeat sequence series (Fig. 4). A summary of the sequence comparison performed within all repeat elements is showed in Table I. The identities were as high as 98% within h3'R_{146} repeats and as low as 56.4% for the 5'R_{120} repeat. Nevertheless, no evident identities were found among the different type of repeat sequences (data not shown).

Upstream of the UPE is the proximal terminator (T₀) that significantly stimulates rRNA transcription. In carp IGS, the sequence GGGTGACCAGTGG, which probably corresponds to the transcriptional terminator, was found upstream of the +1 and downstream of the 28S rRNA. This observation strongly suggests that this element could be the proximal terminator (T₀), which in other species is involved in transcription initiation stimulation (Labhart, 1994, 1995). Interestingly, carp rDNA shows two T₀ elements integrated in a large sequence repeat, and to our knowledge, the carp is the only specie to show this feature, a fact that could represent an additional mechanism for transcription initiation control in eurythermal fish.

It is well known that the nucleolus build-up is organized in the loci containing rDNA arrays (Scheer and Weisenberger, 1994), nevertheless, the mechanisms governing the changes in rRNA expression and their relationship with the morphological cyclic changes of the nucleolus under the acclimatization process remains a major question. The putative cis-elements described here can be used as an important structural basis to future studies on the seasonal transcriptional regulation of the carp rDNA.

Ribosomal DNA (rDNA) is a remarkable case of gene duplication and has been used as a model to test the concerted evolution theory (Hillis and Dixon 1991; Reed et al., 2000). Due to their moderate rate of sequence evolution, coding- and non-coding (ITS1 and ITS2) transcribed sequences of the rDNA are significant tools for phylogenetic and population studies (Reed et al., 2000). Inversely, phylogenetically-supported folding models have been used in fish, including the carp, to discern functional structural features that were preferentially preserved among different species (Joseph et al., 1999).

As expected, 18S rRNA coding sequences comparison reveals a high degree of conservation through the evolution (data not shown). Therefore, an unrooted phylogram was constructed to analyze the relationship of the C. carpio 18S rRNA gene with all other available 18S rRNA sequences from fish, as well as non-related taxa, such as plants, yeast, bacteria, and humans (Table II and Fig. 5

| TABLE I |
|-----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|                | 1      | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      |
| 5'R_{120}      | 100    | 91.8   | 79.8   | 56.4   | 72.4   | 80.2   | 73.0   | 77.8   | 64.7   |
| 5'R_{110}      | 100    | 89.0   | -      | -      | -      | -      | -      | -      | -      |
| 5'R_{330}      | 100    | 91.7   | -      | -      | -      | -      | -      | -      | -      |
| 3'R_{180}      | 100    | 95.0   | -      | -      | -      | -      | -      | -      | -      |
| h3'R_{146}     | 100    | 98.0   | 98.0   | -      | -      | -      | -      | -      | -      |
| 3'R_{146}      | 100    | 98.7   | 71.4   | -      | -      | -      | -      | -      | -      |
The unrooted phylogram shows that fish group in three clades, two of which correspond to the class Actinopterygii and one to the class Elasmobranchii. It is interesting to observe that the different species that constitute each fish clade are sorted according to their order. Furthermore, the distribution of the different species within the cladogram is supported by phylogenetic trees built using different genes (Zardoya and Meyer, 1996; Ishikawa, 2000; Wittbrodt et al., 2002).

ACKNOWLEDGEMENTS

We thank Margaret Snook for critical reading of the manuscript. This work was supported by Grants 1000-061 from FONDECYT and 93-01, 94-01 and 96-01 from the Research Fund of UNAB.

TABLE II

Species and GenBank accession numbers of the full length sequences used for the phylogenetic analysis of the 18S rRNA. (*) Zebrafish 18S rRNA gene sequence was compiled from non-assembled contigs available at the non-processed zebrafish sequence database (http://trace.ensembl.org/perl/ssahaview?server=danio_rerio). (**) Spotted green pufferfish 18S rRNA gene sequence is available under EMBL accession number DS42722.

<table>
<thead>
<tr>
<th>Organism</th>
<th>GenBank Accession Number</th>
<th>Common name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Cyprinus carpio</td>
<td>AF133089</td>
<td>Carp</td>
</tr>
<tr>
<td>2 Danio rerio</td>
<td>*</td>
<td>Zebrafish</td>
</tr>
<tr>
<td>3 Tetraodon nigroviridis</td>
<td>**</td>
<td>Spotted green pufferfish</td>
</tr>
<tr>
<td>4 Xenopus laevis</td>
<td>X02995</td>
<td>African clawed frog</td>
</tr>
<tr>
<td>5 Homo sapiens</td>
<td>X03205</td>
<td>Human</td>
</tr>
<tr>
<td>6 Mus musculus</td>
<td>X00686</td>
<td>House mouse</td>
</tr>
<tr>
<td>7 Arabidopsis thaliana</td>
<td>X16077</td>
<td>Thale cress</td>
</tr>
<tr>
<td>8 Saccharomices cerevisiae</td>
<td>Z75578</td>
<td>Baker’s yeast</td>
</tr>
<tr>
<td>9 Escherichia coli</td>
<td>Z83205</td>
<td>-</td>
</tr>
<tr>
<td>10 Polyodon spathula</td>
<td>AF188371</td>
<td>Mississippi paddlefish</td>
</tr>
<tr>
<td>11 Dalatias licha</td>
<td>AY049827</td>
<td>Kitefin shark</td>
</tr>
<tr>
<td>12 Mitsukurina owstoni</td>
<td>AY049840</td>
<td>Goblin shark</td>
</tr>
<tr>
<td>13 Ictalurus punctatus</td>
<td>AF021880</td>
<td>Channel catfish</td>
</tr>
<tr>
<td>14 Oncorhynchus mykiss</td>
<td>AF308735</td>
<td>Rainbow trout</td>
</tr>
<tr>
<td>15 Squatina californica</td>
<td>AY049858</td>
<td>Pacific angelshark</td>
</tr>
<tr>
<td>16 Pristiophorus cirratus</td>
<td>AY049849</td>
<td>Longnose sawshark</td>
</tr>
<tr>
<td>17 Rhinobatos productus</td>
<td>AY049852</td>
<td>Shovelnose guitarfish</td>
</tr>
<tr>
<td>18 Oryzias latipes</td>
<td>AB105163</td>
<td>Japanese rice fish</td>
</tr>
<tr>
<td>19 Clupea harengus</td>
<td>X98845</td>
<td>Atlantic herring</td>
</tr>
<tr>
<td>20 Megalops atlanticus</td>
<td>X98846</td>
<td>Tarpon</td>
</tr>
<tr>
<td>21 Elops hawaiensis</td>
<td>X98841</td>
<td>Hawaiian ladyfish</td>
</tr>
<tr>
<td>22 Fundulus heteroclitus</td>
<td>M91180</td>
<td>Killifish</td>
</tr>
<tr>
<td>23 Sebastolobus altivelis</td>
<td>M91182</td>
<td>Longspine thornyhead</td>
</tr>
<tr>
<td>24 Squalus acanthias</td>
<td>M91179</td>
<td>Spiny dogfish</td>
</tr>
<tr>
<td>25 Latimeria chalumnae</td>
<td>L11288</td>
<td>Coelacanth</td>
</tr>
<tr>
<td>26 Salmo trutta</td>
<td>X98839</td>
<td>Brown trout</td>
</tr>
</tbody>
</table>
Figure 5
Neighbor-joining phylogenetic tree from full-length 18S rRNA sequences. The bootstrap values are included for each branch.

REFERENCES


