Characterization of repetitive DNA sequences carrying 5S rDNA of the triploid gimbuna (Japanese silver crucian carp, *Carassius auratus langsdorfi*)

Masaru Murakami and Hideo Fujitani*

Laboratory of Molecular Biology, Azabu University School of Veterinary Medicine, 1-17-71 Fuchinobe, Sagami-hara, Kanagawa 229-0006, Japan

(Received 9 October 1997, accepted 20 December 1997)

Repetitive DNA sequences (Hi-b; 209 bp in length) were isolated from the *Hind*III digests of the genomic DNA of the triploid gimbuna, *Carassius auratus langsdorfi*. Sequence analyses revealed that the Hi-b repetitive units were comprised of the complete coding regions of 5S rDNA (120 bp in size) and their 5' flanking regions. The sequences of the Hi-b units from the same individual were highly homogeneous. Southern blot hybridization to the Hi-b probe displayed intricate patterns that represented the presence of other repetitive units containing the Hi-b related sequences. A major family of repetitive sequences related to the Hi-b was then obtained by the polymerase chain reaction using asymmetry primers for the 5S coding regions. These 331-bp sequences (AZ5S's) contained 5S pseudogenes as well as the almost entire Hi-b sequences, and seemed to be the true 5S rDNAs. The tandem arrangements of the AZ5S sequences explained most of the complex results of Southern blots. Another class of intriguing repeat units (Hi-b- β and Hi-b- γ) were also isolated. Fluorescence in situ hybridization data revealed two major signals on a pair of homologous chromosomes and several minor signals on other chromosomes in the triploid gimbuna, indicating the existence of the 5S related sequences as several separate clusters. The major spots were shared with the tetraploid gimbuna and goldfish, but not with the diploid gimbuna. When the genomic organization of the Hi-b related sequences in other cyprinid fishes was examined, the hybridization patterns of the gimbuna were very similar to those of the goldfish, but were clearly different from those of the gengorobuna. The carp genome showed less complex patterns. Thus, the present 5S rDNA-related sequences could be candidates for phylogenetic molecular markers for the crucian carp.

INTRODUCTION

The so-called gimbuna (Japanese silver crucian carp, *Carassius auratus langsdorfi* or *langsdorfii*), which are widely distributed in Japan, occurs in the triploid and rare tetraploid forms in addition to the diploid form (Kobayasi et al., 1970; Liu et al., 1980; Kobayasi, 1982). Remarkably, the polyploid gimbunas (triploids and some tetraploids) reproduce by means of the natural gynogenesis, resulting in clonal, all-female offspring (Kobayasi, 1971; Kobayasi and Ochi, 1972; Ojima and Asano, 1977). Such an unusual vertebrate should provide a good system for the study of clonal reproduction and for the elucidation of the usual reproductive mechanisms in vertebrates.

However, the genetic background of the gimbuna, such as the phylogenetic relationship among subspecies of *C. auratus*

and the origin of the polyploid gimbuna, is not well understood, though some cytological or isozyme-electrophoretic studies have been reported (Ojima et al., 1975; Kobayasi and Nakano, 1976; Shimizu et al., 1993). The paucity of morphologically distinguishing traits makes it difficult to precisely identify the gimbuna within the species *C. auratus*. Thus, careful molecular analyses of the genome are required not only for the elucidation of the evolutionary trends of the gimbuna but also for the resolution of taxonomic confusion regarding *C. auratus*.

Eukaryotic genomes contain a number of families of repeated DNA sequences, including satellite DNAs (Brutlag, 1980; Singer, 1982), multigene families such as rDNAs, LINEs and SINEs (Weiner et al., 1986). Since these repeated DNA sequences show both high intraspecific or intrapopulation homogeneity and high interspecific or interpopulation variability, they have been considered as informative genetic markers for the identification of a species

* Corresponding author.

or a population. They are used also for the determination of the phylogenetic relationship among closer species in a wide variety of organisms including various vertebrates (Quinn et al., 1992; Murata et al., 1993; Wijers et al., 1993; Mizuno et al., 1995; Takasaki et al., 1996). However, not a great deal is known about repetitive DNA sequences in the genome of cyprinid fishes at present. We therefore are focusing on repetitive DNA sequences of the gimbuna for a better understanding of its phylogenetic situation. We so far found two intriguing families of satellite DNA sequences from the gimbuna genome; one was characteristic of the polyploid gimbuna specific sequences (Murakami and Fujitani, 1997a) and the other was common to *C. auratus* populations (Murakami and Fujitani, 1997b). This paper reports another characteristic repetitive DNA sequences from the gimbuna, which contains the putative 5S rRNA gene known as one of the multigene families. Our efforts were focused mainly on characterization and structural analysis of the 5S rDNA-related sequences in the genome of the triploid gimbuna. To our knowledge, this is the first description with regard to the 5S rDNA repetitive sequences of the *Carassius* species or genus.

MATERIALS AND METHODS

Fishes and genomic DNA samples. Gimbunas and common carp (*Cyprinus carpio*) were collected from the Shibuta River, Kanagawa, Japan. A clonal line of triploid gimbuna (AZ3 line), which originally taken from the same river, has been gynogenetically reared up to three generations in this laboratory. Goldfish (*C. a. auratus*) and a commercially cultured form of gengorobuna (*C. a. cuvieri*) were purchased from local fish dealers. The ploidy of gimbunas was determined by calculating the ratio of fluorointensity of their erythrocyte DNA to that of the goldfish sperm DNA with DAPI staining described by Hamada and Fujita (1983) with minor modifications.

Genomic DNA was isolated from the liver tissue according to the standard proteinase K/SDS digestion and phenol extraction method (Sambrook et al., 1989).

Isolation and sequencing of the repetitive DNAs. The genomic DNA from a triploid gimbuna (AZ3 line) or a goldfish was completely digested with *Hind*III. The digests were fractionated on a preparative 3.0% agarose gel and were visualized with ethidium bromide. A fragment of 210 bp in length was excised and ligated into the *Hind*III-linearized plasmid vector pUC19. The ligated DNAs were transformed into competent *E. coli* JM109. In some cases, fragments of about 300 and 500 bp in size were cloned. Transformants were selected on LB agar plates containing ampicillin, X-gal, and IPTG, and recombinant plasmids were confirmed by size. The WizardTM Minipreps Resin (Promega) was used for the purification of the plasmids. Both strands of each plasmid were sequenced with an

AmpliTaQ Dye Terminator Cycle Sequencing FS kit (Perkin-Elmer) using an 373A autosequencer (Applied Biosystems).

Polymerase chain reaction (PCR) amplification, cloning and sequencing. The PCRs were performed with 100 ng of genomic template DNA, 200 μ M each of dNTPs, 2.5 U *Taq* polymerase (TaKaRa) and 1 μ M each of primers: 5' TATGCCCGATCTCGTCTGATC 3' and 5' CAGGTTGGTATGGCCGTAAGC 3', which were designed based on the sequence of the putative 5S coding region of the AZHi-b2 (see Fig. 2). The cycling conditions were as follows: initial denaturation at 95°C for 4 min, 30 cycles of annealing at 63°C for 30 sec, extension at 72°C for 1 min, denaturation at 95°C for 1 min and then a 9 min extension at 72°C. The products were electrophoresed and stained with ethidium bromide. An about 330-bp band of the products from the AZ3 gimbuna was cut out of the gel and reamplified under the same PCR conditions as described above except that 0.8 ng of the template was used. The reamplified products were purified by a MicroconTM 50 (Amicon) to remove the excess of primers and were verified to be approximately 330 bp long by agarose gel electrophoresis. After the purified product was treated with T4 DNA polymerase (TaKaRa) and T4 polynucleotide kinase (GIBCO-BRL), it was ligated into the *Sma*I site of the pUC19. Recombinant clones were selected and the inserts were sequenced as mentioned above.

Southern blot hybridization. Five microgram of genomic DNA samples were separately cleaved to completion with *Hind*III, *Dra*I, *Fok*I, *Sca*I or *Eae*I at 8 U/ μ g DNA, fractionated on 2.0% agarose gels and transferred to MagnaGraph nylon membranes (Micron Separations Inc.). Biotin-labeling of the probe and hybridization were performed as described previously (Murakami and Fujitani, 1997a). The labeled CaHi-b3 (see Fig. 2) was employed as a probe in this study. Post-hybridization washes were carried out at high stringent conditions: washes with a solution of 0.16 $\times$ SSC and 0.1% SDS at 65°C. Chemiluminescent hybridization signals were detected with PhototopeTM-Star detection kit (New England Biolab) according to the manufacturer's instructions.

Fluorescence in situ hybridization (FISH). Metaphase chromosomes were obtained directly from the kidney tissue of the fish intraperitoneally colchicized for several hours. The slides were hardened at 65°C for 4 h and were treated with 100 μ g/ml RNase (Sigma Chemical) prior to hybridization. For each sample, 500 ng of the whole plasmid (CaHi-b3, see Fig. 2) was nick-translated using a premixed solution with biotin-16-dUTP (Boehringer Mannheim) according to the manufacturer's instructions. The biotinylated probe was ethanol-precipitated with 5 μ g each of sonicated salmon sperm DNA (Sigma Chemical) and *E. coli* tRNA (Sigma Chemical), and heat-denatured in formamide.

Hybridization procedure was as described previously (Murakami and Fujitani, 1997a) except that the above probe was used. Signals were detected using 5 µg/ml avidine-FITC (Boehringer Mannheim). Chromosomes were counterstained with 1 µg/ml propidium iodide (PI, Sigma). Samples were visualized with fluorescence microscope (BX-FLA, Olympus) using standard filter sets for FITC and PI. At least five metaphase spreads per specimen were examined.

RESULTS

Cloning and sequencing of repetitive DNAs. When the genomic DNA of the triploid gimbuna, diploid gimbuna, goldfish or gengorobuna was digested with *Hind*III and the fragments electrophoresed, discrete bands of about 210 bp and 270 bp in length were observed (Fig. 1). The 270-bp fragment was dealt with elsewhere (Murakami and Fujitani, 1997b). The other prominent fragment of 210 bp was cloned into the plasmid vector pUC19. The recombinants were identified by size. The complete nucleotide sequences of the inserts, which we designated as Hi-b's, from a triploid gimbuna (AZ3 clonal line) and a goldfish were determined (Fig. 2). Although all Hi-b sequences were unique, the nucleotide similarity among the AZHi-b (Hi-b from the

AZ3 fish) monomers and between the AZHi-b consensus (AZHi-b2) and the CaHi-b3 from the goldfish were as high as 95.8% on the average and 97.6%, respectively. Homology search with the DDBJ DNA database revealed that the sequences from position 90 to the 3' end of Hi-b repetitive units (120 bp long) were very similar to the entire 5S-rRNA gene reported for other fishes (more than 91% identity, Fig. 2). In addition, the Hi-b sequences, especially the putative 5S coding regions, were similar to the tandemly repeated sequences (186 to 201 bp long) from three cyprinid species, namely, the silver carp, bighead carp, and grass carp (Huang et al., 1993). If Hi-b sequences were divided into two regions, the putative 5S rDNA coding regions of Hi-b2, 6, 7 and 11, and the flanking regions of Hi-b8 and 13 or those of Hi-b12 and CaHi-b3 were identical. Point mutations appeared to occur evenly throughout the entire repeat unit; 22 (18.3%) and 20 (22.5%) of nucleotide positions were variable in the putative 5S coding regions and the flanking regions, respectively. A clone with the insertion of clusters of A's and T's (a total of 13 bp) at the boundary of the putative 5S coding region and the flanking region was also isolated (data not shown), indicating the presence of a 222-bp variant of Hi-b repetitive sequences.

Southern blot hybridization analyses. In order to examine the genomic organization of Hi-b repetitive DNA sequences in the triploid gimbuna, the genomic DNA was digested with *Hind*III and hybridized to a Hi-b probe (Fig. 3, lane 3n), to begin with. The digests exhibited three strongly hybridizing bands of about 210, 340, and 500 bp, as well as several minor bands of larger size. Apparently the 340- and 500-bp bands were not the oligomers of the Hi-b repetitive monomer (209 bp). This result indicates the existence of some larger repetitive units containing the Hi-b sequences or the sequences similar to the Hi-b. Next, the genomic DNA of the triploid gimbuna was treated with each one of *Fok*I, *Dra*I, *Eae*I, and *Sca*I (Fig. 3, lane 3n), since almost all of the Hi-b units sequenced here had a single recognition site for each of them (Fig. 2). Most of the *Fok*I digests gave three hybridizing bands, that is, a major band of 340 bp and two minor bands of 200 bp and 500 bp, which were similar in size to those of *Hind*III digests. In *Dra*I digests, a 210-bp band of the Hi-b repeat monomer was hardly discernible, and instead, an intense band of 280 bp as well as 500 bp was observed. These results argue that the Hi-b sequences are included in the repeat units of larger size, and one of the dominant repeat units seems to be 340 bp in size. *Eae*I digests yielded highly intense bands with high molecular weights and several less intense bands of 500, 640, 780 bp, etc. with about 140 bp increment, but no 210-bp band of Hi-b monomer or 340-bp band. *Sca*I digests exhibited a number of bands encompassing a wide size range, which indicated the presence of the nonperiodic and multiple *Sca*I sites within the clusters of the repetitive units. These hybridization pat-

Fig. 1. Detection of repetitive DNA sequences in *Hind*III digests of the genomic DNAs from the triploid gimbuna (3n, AZ3 line), diploid gimbuna (2n) and goldfish (GF) on a 3.0% agarose gel. The arrow designates a satellite band of approximately 210 bp. The size-marker lane (M) is a 100-bp ladder. The gengorobuna exhibited the same electrophoretic pattern as the gimbuna did.

Fig. 2. Nucleotide sequences of Hi-b monomers. Eleven randomly selected clones from a triploid ginbuna (AZ3 line) and one clone (CaHi-b3) from a goldfish were sequenced and aligned using a Clustal W program. Only the complete sequence of AZHi-b2, which is identical with the consensus sequence determined by the majority rule, is shown and the sequence differences relative to it are indicated. Dots and a dash signify identical nucleotides and a sequence gap, respectively. Substitutions are marked with the appropriate nucleotides. The putative 5S coding region is underlined, and the 5S rRNA genes of some other teleost fishes are referred to. The abbreviations of the fishes are as follows; Tt: *Tinca tinca*, Tench (Dennis and Wegnez, 1977), Sa: *Salmo salar*, Atlantic salmon, Sg: *Salmo gairdnerii*, rainbow trout, and Mf: *Misgurnus fossilis*, loach. Each accession number is indicated in parentheses. The recognition sites for five restriction enzymes (*Hind*III: AAGCTT, *Dra*I: TTTAAA, *Eae*I: YGGCCR, *Sca*I: AGTACT, *Fok*I: GGATG) are shown above the sequences. Some monomers have no *Sca*I and/or *Eae*I sites. The pair of asymmetric primers for the PCR experiment (Fig. 4) are indicated with arrows. All sequences reported here are available from the DDBJ/EMBL/GenBank databases under accession Nos. AB001492–001503.

examined, although some of the diploid ginbuna showed slight differences in size or intensity of certain bands.

Fig. 3. Southern blots of restriction digests of the genomic DNAs from the triploid ginbuna (3n), diploid ginbuna (2n), goldfish (GF), gengorobuna (GRB) and common carp (CC) with the Hi-b probe. The numbers show the size of markers in bp.

Isolation of larger repetitive units with sequences related to Hi-b. A PCR amplification technique was

employed to obtain the predicted larger repetitive units containing the Hi-b by using asymmetry primers for the

putative 5S coding region, since the 5S rDNA is known to be organized into tandem and head to tail arrangement in higher eukaryotes. The primers were designed according to Pendas et al. (1994). In our PCR (Fig. 4), there were several electrophoretic bands of the amplified fragments in addition to a prominent fragment of about 330 bp as expected from the Southern blot data. These fragments exhibited slight variation in size in some individuals of the triploid gimbuna. The dominant 330-bp fragment from a gimbuna (AZ3 line) was cloned into pUC19. The inserts of nine clones, collectively designated as AZ5S, were sequenced (Fig. 5).

All the AZ5S sequences were 331 bp in size and comprised the whole Hi-b monomer but with a deletion of 7 bp, and an interposed region (IPR) of 137 bp between adjacent Hi-b monomers. The most remarkable feature of this IPR (position 98 to 234) was that it contained a pseudogene of the 5S rRNA, which was the truncated 5S gene corresponding to 79 nucleotides of its 5S coding region lacking the first 41 nucleotides from the 5'-terminus, with 89.9% homology. The first 76-bp stretch (position 235 to 310) of the Hi-b-equivalent region bore no significant resemblance to the first 58 bp (position 98 to 155) of the IPR. Clusters of T's and a direct repeat (TATATAATG) were located closely downstream of both the 5S coding region and its pseudogene region. The 5S coding regions of the AZ5S-a7, a1, f6, f8,

and e6 were identical to those of the Hi-b-2, 6, 7 and 11. The total non-coding region (211 bp) excluding the 5S coding region was AT-rich (61.6%) with some clusters of A or T residues. The promoter sequences for polymerase III, termed box A and box C (Ciliberto et al., 1983), were present in the putative 5S coding region. Moreover, clusters of five consecutive T residues were seen at 3' boundary of the coding region, which could serve as a polymerase III termination site (Bogenhagen and Brown, 1981). None of these conserved sequences were modified by mutations.

Nine clones could be divided into two groups on the basis of the nucleotide sequence; one group, termed group I, was composed of the clones a7, d2, a1, a3, f6, f8, and d1 having a single *Hind*III recognition site, and the other, termed group II, b6 and e6 with two *Hind*III sites. The extent of sequence variation within each group was less than 1.2%, while that between the group I and group II was more than 4%. The group II, determined by digestion of inserts with *Hind*III, accounted for 12% of the 42 positive clones.

We also searched for larger repetitive units in the *Hind*III digests of the triploid gimbuna (AZ3 line) genome. Although we failed to isolate repetitive DNA sequences containing the Hi-b as in the case of AZ5S, other fascinating repetitive DNA sequences were obtained. As shown in Figure 6, these sequences consisted of a duplicate, named Hi-b- β , or a triplicate, named Hi-b- γ , of subrepeats (Hi-b- α : approximately 140 bp in length), which corresponded to the IPR in the AZ5S.

Fig. 4. Electrophoresis (1.8% agarose gel) of the asymmetrically primed PCR products from the genomic DNAs of the triploid gimbuna (3n), goldfish (GF), and gengorobuna (GRB). A predominant band (about 330 bp), which was seen in the triploid gimbuna and the goldfish, is arrowed. The size-marker lane (M) is a 100-bp ladder.

Genomic organization of Hi-b sequences in other cyprinid fishes.

By using a Hi-b sequence as the probe for Southern hybridization, we examined the genomes of other cyprinid fishes; three goldfish (*C. a. auratus*), five gengorobuna (*C. a. cuvieri*), and four common carp (*Cy. carpio*) (Fig. 3). The hybridization patterns were essentially identical among individuals of the same (sub)species. The patterns of the goldfish were considerably similar to those of the triploid gimbuna with all restriction enzymes used, while those of the gengorobuna differed from those of the gimbuna or goldfish. Particularly striking differences were observed in *Dra*I digests of the gengorobuna genome. There were strongly hybridized bands of 210 bp and 420 bp, which were practically absent from the *Dra*I-digested gimbuna genome. The 280-bp band was shared by the both parties. The gengorobuna's 210- and 420-bp bands were seen as the major bands also in both the *Hind*III- and *Fok*I-digests and coincided with the monomer and dimer of the Hi-b repeat unit, respectively. The patterns for the carp genome digested with any of the enzymes except *Eae*I were very simple; they showed only a major 210-bp band of the Hi-b monomer and its oligomeric bands with faint signals, representing the tandem arrangement.

FISH analyses. Eight individuals of the triploid gimbuna including one specimen of the AZ3, six individuals of the

I

AZ5S-a7 1: TATGCCCGAT CTCGTCTGAT CTCGGAAGCT AAGCAGGTTT GGCCTGGTT AGTACTTGGA TGGGAGACCG
AZ5S-d2 :
AZ5S-a1 :
AZ5S-a3 :
AZ5S-f6 :
AZ5S-f8 :
AZ5S-d1 :G.
AZ5S-b6 :
AZ5S-e6 :

II

AZ5S-a7 71: CCTGGGAATA CCAGGTGCTG TAAGCTTTTT GGAAATTTTT CTCTTAGTAT ATAATGATT TGCCAAAAA
AZ5S-d2 :T.....
AZ5S-a1 :
AZ5S-a3 :T..
AZ5S-f6 :
AZ5S-f8 :G.
AZ5S-d1 :
AZ5S-b6 :T.....C...A...G..
AZ5S-e6 :

HindIII

((ScaI))((FokI))

AZ5S-a7 141: TAGAGTCAAT GCCCAATCTC TGAATATTAG CAGGTTTGGG CCTGGTGTAG ACATGGATGG GAGACTGCCT
AZ5S-d2 :
AZ5S-a1 :
AZ5S-a3 :
AZ5S-f6 :
AZ5S-f8 :
AZ5S-d1 :
AZ5S-b6 :G....C....A..T.....T.T...A..
AZ5S-e6 :G....C....A..T.....T.T...A..

(DraI)((HindIII))

AZ5S-a7 211: GGGAATACCA GGTGCTTTAA ACTTATCTTT CCTACTTATA TAATGTACTG GCGATTAGAT TGGCTGGTCT
AZ5S-d2 :
AZ5S-a1 :T.....
AZ5S-a3 :
AZ5S-f6 :
AZ5S-f8 :T.....
AZ5S-d1 :
AZ5S-b6 : :ATT.....G...T.....
AZ5S-e6 : :ATT.....G...T.....

DraI EaeI

AZ5S-a7 281: TTAATAGCC CTCTCTTTGC AGCAGTCTTC GCTTACGGCC ATACCAACCT G 331
AZ5S-d2 :
AZ5S-a1 :A.....
AZ5S-a3 :
AZ5S-f6 :C.....
AZ5S-f8 :
AZ5S-d1 :
AZ5S-b6 :
AZ5S-e6 :

Fig. 5. Nucleotide sequences of AZ5S monomers. Base changes with regard to the reference sequence of the AZ5S-a7 are shown for each monomer. Two groups (type I and type II) of monomers can be easily distinguished by several distinct sites of mutation. Five kinds of restriction sites are shown above the sequences, and those that are present only in type I sequences are indicated in parentheses. The 2nd *ScaI* and *HindIII* sites bracketed with double parentheses occurs only in the type II sequences. The internal control regions for polymerase III, referred to as box A and box C, exist within the putative 5S coding region (underlined). The 5S pseudogene (position 156 to 234) is boxed. The dashed arrows denote a direct repeat. The inverted PCR primers are indicated by the solid arrows. A total of 200 bp extending from position 229 to 331 and from position 1 to 97, except for a 7-bp deletion (GGGTTT) between position 236 and 237 and for a lack of 2 bp (GC) caused by the design of PCR primers, corresponds to the Hi-b sequences. All sequences reported here have been submitted to the DDBJ and assigned accession Nos. AB007776–007784.

AZ5S-a7- IPR	98:	C.....T.....T.....A.....A	
Hi-b-γ1	1:	AAGCTTTTGGAAATTTTATCTTAGTACATAATGAATTTGCCAAGAAATAGAGTCAATGCCCG	
Hi-b-γ6	1:		
Hi-b-β4	1:	T.....CA.....T.....A.T.....A.....	
Hi-b-β5	1:	C.....T.....A.T.....TA.....	
Hi-b-β10	1:	C.....CA.....T.....A.T.....A.....	
Hi-b-β13	1:	C.....CA.....T.....A.T.....A.....	
AZ5S-a7- IPR	156:		234
Hi-b-γ1	65:	ATCTCTGAATATTAGCAGGTTGGGCCTGGTTAGTACATGGATGGGAGACTGCCTGGGAATACCAGGTGCTTTAAACTT	
Hi-b-γ6	65:		
Hi-b-β4	65:		
Hi-b-β5	65:	-.....A.	
Hi-b-β10	65:	T.....T.....T.....	
Hi-b-β13	65:	T..T.....T.....	
AZ5S-a7- IPR	98:	..T....TT...T.TT.G.....G..T..GCC.....T.-----.....A	
Hi-b-γ1	144:	TTTCGAAA-ATTTACAACTATATAATAATCTTAAAAAAAAAAAAAAAAAAAAA---GAGTCAATGCCCG	
Hi-b-γ6	144:		
Hi-b-β4	144:	..T.....-.....T.....-----.....A	
Hi-b-β5	143:	..T.....-.....T.....-----.....	
Hi-b-β10	144:	..T.....-.....G..T.....TC.TT.....AAAA.....	
Hi-b-β13	144:	..T.....-.....G..T.....TC.TT.....-----.....	
AZ5S-a7- IPR	156:		234
Hi-b-γ1	215:	ATCTCTGAATATTAGCAGGTTGGGCCTGGTTAGTACATGGATGGGAGACTGCCTGGGAATACCAGGTGCTTTAAACTT	
Hi-b-γ6	215:		
Hi-b-β4	210:	T..C.....C..A..T.....T.....A....C..G..A.....G.....G...	288
Hi-b-β5	209:	C.....A..T.....T.T....A.....A.....G...	287
Hi-b-β10	219:	C.....A..T.....T.T....A.....A.....G..G...	297
Hi-b-β13	213:	C.....A..T.....T.T....A.....A.....G.....	291
AZ5S-a7- IPR	98:	TT...T.TT.G.....G..T....C.....T.-----.....A	
Hi-b-γ1	294:	TTTGGAAA-ATTTACAACTATATAATAATCTTGCAAAAAAAAAAAAAAAAAAAAAA---GAGTCAATGCCCG	
Hi-b-γ6	294:	A.....	
AZ5S-a7- IPR	156:	A.....G..C.....A.G....G.....G.....A...	234
Hi-b-γ1	367:	ATCTCTGAATCTTAGCAGGTTAGGTCTGGTTAGTACTTTGATGAGAGACTGCCTAGGAATACCAGGTGCTTTAAGCTT	445
Hi-b-γ6	368:		446

Fig. 6. Sequence alignment of Hi-b-β and Hi-b-γ monomers, and sequence comparison of those and the interposed region (IPR) within the AZ5S. Four clones of Hi-b-β and two clones of Hi-b-γ were sequenced. Alignments were performed using Clustal W and with eyes. Nucleotide differences relative to the complete sequence of Hi-b-γ1 are shown. Dots and dashes indicate identical nucleotides and gaps, respectively. The 5S pseudogene in the IPR is boxed. The Hi-b-β and Hi-b-γ are constituted of two and three tandem subrepeats, respectively, and each subrepeat (approximately 140 bp in length), except for insertions of a stretch of A's in the 2nd subrepeat of the Hi-b-β and Hi-b-γ and in the 3rd subrepeat of the Hi-b-γ, corresponded to the IPR in the AZ5S sequences (position 98 to 234, see Fig. 5.). The Hi-b-β and Hi-b-γ sequences are available from the DDBJ/EMBL/GenBank DNA databases under accession Nos. AB001504–001507 and AB007785–007786, respectively.

diploid ginbuna, one female tetraploid ginbuna and five goldfish (Wakin, two females and three males) were examined for the presence of the hybridization signals with the Hi-b probe. In all the triploid ginbuna, two major signals were located on two chromosomes, which seemed to be a pair of large-sized submetacentric chromosomes as identified by PI counterstaining (Fig. 7A). Computer image analyses allowed us to assign these two signals to the short arms of a large submetacentric pair. Six to ten additional minor signals (some clear, others less clear) were observed on separate chromosomes. This observation indicates that

the 5S rDNA-related sequences are localized not only as a large cluster but also as several clusters in the triploid ginbuna, and this fact justifies the complex structure of the 5S related sequences in the triploid ginbuna genome. It was impossible to allocate these minor signals on particular chromosomal loci because of the large number (about 156) and small size of the chromosomes. The relationship between the fluorescent signal on each chromosome and certain particular 5S rDNA-related sequences could not be determined in the present study either.

As seen in Figure 7B, the diploid ginbuna had no major

Fig. 7. FISH with the Hi-b probe to metaphase chromosomes from the triploid ginbuna (3n), diploid ginbuna (2n), tetraploid ginbuna (4n), and the goldfish (GF). In A, a representative spread (left side) and its partial enlargement (right side) from a triploid ginbuna are shown. Large and small arrowheads designate the location of the major and minor hybridization signals, respectively. Two chromosomes with the major signals are boxed. In B, only major signals in each fish are marked by large arrowheads. PI counterstaining of the same spreads is also presented.

signals but only several minor ones (some of them clear, others less clear). By contrast, the 205 metaphase chromosomes from a tetraploid gibel carp exhibited four strong signals on separate chromosomes as well as several weak signals on other chromosomes. In the female goldfish (Wakin), two major signals corresponding to those in the triploid gibel carp were seen whereas there was only one in the male.

DISCUSSION

We have isolated, from the triploid gibel carp genome, the 5S rDNA-related repetitive DNA sequences: the Hi-b, AZ5S, Hi-b- β , and Hi-b- γ . A partial genomic organization of these repetitive DNA sequences, which were inferred from the sequences and Southern blot data, are schematically presented in Figure 8. Such a side-by-side arrangement of the 5S rRNA gene and its pseudogene has been reported also in the 5S rDNA of *Xenopus* genome (Miller and Brownlee, 1978; Miller et al., 1978). The sequence analysis data suggested the AZ5S's are most likely the true 5S rDNA sequences. Clusters of T's and a direct repeat in vicinities of both ends of the IPR might have played a role in the insertion/deletion of the IPR with the pseudogene into, or from, the AZ5S. The AZ5S monomers with a few nucleotide substitutions in the 5S coding region are possibly the variants or oocyte- or somatic-type of 5S rRNA gene, as found in other teleosts (Dennis and Wegnez, 1977; Mashkova et al., 1981; Sedman et al., 1989), although we cannot exclude the possibility of the *Taq* polymerase misincorporation during the course of reamplification. The amplification primed by the two inverted primers designed to anneal within the 5S coding region assures the tandem connection of at least two AZ5S units. The tandem arrangement of the type I sequences can explain the strong hybridizing bands of the 340 bp in the *Hind*III digests, of the 280 bp in the *Dra*I digests, and of the 200 bp in the *Fok*I digests observed by Southern blots. Similarly, the

intense hybridization signals of the 210-bp band in the *Hind*III digests and 340-bp band in the *Fok*I digests in the Southern blot patterns would have come from the tandem arrangements of the type II sequences. The tandem repeats of AZ5S units thus seemed to be organized mainly in the triploid gibel carp genome. Despite the presence of an *Eae*I site in all of the AZ5S's sequenced here, the hybridization results with *Eae*I digests showed no 340-bp band corresponding to an AZ5S. This indicates that most of the CpG dinucleotides of the *Eae*I recognition sites in the 5S coding region of the AZ5S are methylated. A partial methylation of the 5S rDNA sequences have been known for *X. laevis* (Miller et al., 1978).

Possible insertions of the Hi-b- β or Hi-b- γ , which comprise the tandem repeats of the IPR, into the AZ5S most likely account for the larger size of minor bands (209 + 290 and 209 + 430 bp in length). We consider the sequences of the Hi-b- β and Hi-b- γ are worth taking note. We would pay attention also to the composite-structure consisting of three regions; a 5S-related region (pseudogene) that contains slightly altered internal promoters of box A and box C for polymerase III, a highly conserved stretch of 13 nucleotides (GAGTCAATGCCCG) immediately preceding the pseudogene, and a highly AT-rich region ending with poly A. These sequence features are suggestive of a SINE originated from 5S rRNA, although most SINEs are known to be derived from tRNAs (Okada, 1991). Further careful investigations are required to see if these are one of the SINEs (Endoh and Okada, 1986).

The 5S rDNA sequences have been used as good markers in phylogenetic analysis among closely related species in some vertebrates due to the concerted evolution (Suzuki et al., 1994; Pendas et al., 1995). We therefore compared the genomic structures of the Hi-b sequences in the genomes of several cyprinid fishes. The Southern blot banding patterns of the triploid gibel carp were almost identical to those of the goldfish, but were apparently different from those of

Fig. 8. Inferred schematic representation of the genomic organization of the 5S rDNA-related repetitive sequences (Hi-b, Hi-b- β , Hi-b- γ , and AZ5S) in the triploid gibel carp. The arrows indicate the location of the primers employed in the PCR.

the gengorobuna. Accordingly, the triploid gimbuna appears to be more closely related to the goldfish than to the gengorobuna. The obviously different hybridizing patterns of *Dra*I digests between the gimbuna and the gengorobuna could also be used as a taxonomic marker for analysis of the relationship between these two populations. Sequence comparison of the 5S rDNA-related sequences among those fishes is in progress.

The hybridization patterns of the carp, which belonged to the genus other than the gimbuna, were very simple when compared with those of the gimbuna, goldfish, or gengorobuna. Our speculation is that the complex 5S repeats have evolved after the divergence into the genus *Carassius* or after the subsequent speciation to *C. auratus*, with probable insertion of the pseudogenes between the adjacent Hi-b repetitive units.

Our FISH data showed two major signals on a single chromosome pair in the triploid gimbuna and four in the female tetraploid gimbuna. By contrast, there was no major signal in the diploid gimbuna. This observation and our previous results (Murakami and Fujitani, 1997a) let us harbor doubts about the current classification that the diploid and polyploid (hybrid?) gimbunas belong to the same subspecies, although they are indistinguishable from one another in morphological aspects. Interestingly, the characteristic spots were seen also in the goldfish, in which the female had two whereas the male had one. With regard to these signals, the relationship between the polyploid gimbunas and the goldfish remains to be clarified. In this connection, the chromosomes carrying these major signals seemed to correspond to the second largest submetacentrics with intense C-bands previously described as being markers for sex discrimination in the goldfish (Ojima et al., 1979). It is intriguing to investigate the relationship between the sex chromosome and the 5S rDNA, as reported in the rainbow trout (Moran et al., 1996).

As far as we know, the 5S rDNA of the species *C. auratus* was not identified before. Further analyses of the 5S rDNA-related sequences in the various (sub)species of the genus *Carassius* will contribute to the identification of more precise phylogenetic or taxonomic characters of fishes belonging to the genus *Carassius* of the family cyprinidae and the elucidation of the origin of the gynogenetic polyploid gimbuna.

We are grateful to M. Unoki for her assistance in FISH experiments. This study was supported in part by a Grant-in-Aid for the Encouragement of Young Scientists and a Grant-in-Aid for Exploratory Research from the Ministry of Education, Science, Sports, and Culture of Japan.

REFERENCES

- Bogenhagen, D. F., and Brown, D. D. (1981) Nucleotide sequences in *Xenopus* 5S DNA required for transcription termination. *Cell* **24**, 261–270.
- Brutlag, D. L. (1980) Molecular arrangement and evolution of heterochromatic DNA. *Annu. Rev. Genet.* **14**, 121–144.
- Ciliberto, G., Raugeri, G., Costanzo, F., Dente, L., and Cortese, R. (1983) Common and interchangeable elements in the promoters of genes transcribed by RNA polymerase III. *Cell* **32**, 725–733.
- Dennis, H., and Wegnez, M. (1977) Oocytes and liver cells of the teleost fish *Tinca tinca* contain different kinds of 5S RNA. *Dev. Biol.* **59**, 228–236.
- Endoh, H., and Okada, N. (1986) Total DNA transcription in vitro: a procedure to detect highly repetitive and transcribable sequences with tRNA-like structure. *Proc. Natl. Acad. Sci. USA* **83**, 251–255.
- Hamada, S., and Fujita, S. (1983) DAPI staining improved for quantitative cytofluorometry. *Histochemistry* **79**, 219–226.
- Huang, C. J., Huang, F. L., Wang, Y. C., Hsiao, P. W., and Lo, T. B. (1993) Molecular cloning and relationship of highly repetitive *Hind*III sequences in three cyprinid species: silver carp, big-head carp and grass carp. *Proc. Natl. Sci. Counc. Repub. China Part B Life Sci.* **17**, 85–90.
- Kobayasi, H. (1971) A cytological study on gynogenesis of the triploid gimbuna (*Carassius auratus langsdorffii*). *Zool. Mag. (Tokyo)* **80**, 316–322.
- Kobayasi, H. (1982) The distribution of polyploid funa (*Carassius auratus*) in and around. *Jpn. Women's Univ. J.* **29**, 145–161.
- Kobayasi, H., Kawashima, Y., and Takeuchi, N. (1970) Comparative chromosome studies in the genus *Carassius*, especially with a finding of polyploidy in the gimbuna (*C. auratus langsdorffii*). *Jpn. J. Ichthyol.* **17**, 153–160.
- Kobayasi, H., and Nakano, K. (1976) Morphological and cytological studies in back-cross hybrids of F₁ fishes between the gimbuna (*Carassius auratus langsdorffii*) and the carp (*Cyprinus carpio*). *Jpn. Women's Univ. J.* **23**, 113–122.
- Kobayasi, H., and Ochi, H. (1972) Chromosome studies of the hybrids, gimbuna (*Carassius auratus langsdorffii*) x kinbuna (*C. auratus subsp.*) and gimbuna x loach (*Misgurnus anguillicaudatus*). *Zool. Mag. (Tokyo)* **81**, 67–71.
- Liu, S., Sezaki, K., Hashimoto, K., and Nakamura, M. (1980) Distribution of polyploids of “gimbuna” *Carassius auratus langsdorffii* in Japan. *Bull. Jpn. Soc. Sci. Fish.* **46**, 413–418.
- Mashkova, T. D., Serenkova, T. I., Mazo, A. M., Avdonina, T. A., Timofeyeva, M. Y., and Kisselev, L. L. (1981) The primary structure of oocyte and somatic 5S rRNAs from the loach *Misgurnus fossilis*. *Nucleic Acids Res.* **9**, 2141–2151.
- Miller, J. R., and Brownlee, G. G. (1978) Is there a correction mechanism in the 5S multigene system? *Nature* **275**, 556–558.
- Miller, J. R., Cartwright, E. M., Brownlee, G. G., Fedoroff, N. V., and Brown D. D. (1978) The nucleotide sequence of oocyte 5S DNA in *Xenopus laevis*. II. The GC-rich region. *Cell* **13**, 717–725.
- Mizuno, S., Tajima, S., Saitoh, Y., Mori, K., Ono, H., Tone, M., and Mihira, T. (1995) Variation of repetitive DNA and its phylogenetic relation in Hynobiidae (Caudata). *J. Hered.* **86**, 114–120.
- Moran, P., Martinez J. L., Garcia-Vazquez, E., and Pendas, A. M. (1996) Sex chromosome linkage of 5S rDNA in rainbow trout (*Oncorhynchus mykiss*) Cytogenet. Cell Genet. **75**, 145–150.
- Murakami, M., and Fujitani, H. (1997a) Polyploid-specific repetitive DNA sequences from triploid gimbuna (Japanese silver crucian carp, *Carassius auratus langsdorffii*) Genes Genet. Syst. **72**, 107–113.
- Murakami, M., and Fujitani, H. (1997b) A family of highly repetitive DNAs from “gimbuna” (*Carassius auratus langsdorffii*) genome common to *Carassius auratus* populations. *Zool. Sci. (Tokyo)* **14**, 763–769.
- Murata, S., Takasaki, N., Saitoh, M., and Okada, N. (1993) Determination of the phylogenetic relationships among Pacific

- salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. *Proc. Natl. Acad. Sci. USA* **90**, 6995–6999.
- Ojima, Y., and Asano, M. (1977) A cytological evidence for gynogenetic development of the ginbuna (*Carassius auratus langsdorfii*). *Proc. Jpn Acad. Ser. B Phys. Biol. Sci.* **53**, 138–142.
- Ojima, Y., Hayashi, M., and Ueno, K. (1975) Triploidy appeared in the back-cross offspring from funa-carp crossing. *Proc. Jpn Acad.* **51**, 702–706.
- Ojima, Y., Ueda T., and Narikawa, T. (1979) A cytogenetic assessment on the origin of the gold-fish. *Proc. Jpn Acad. Ser. B Phys. Biol. Sci.* **55**, 58–63.
- Okada, N. (1991) SINEs: short interspersed repeated elements of the eukaryotic genome. *Tree* **6**, 358–361.
- Pendas, A. M., Moran, P., Freije, J. P., and Garcia-Vazquez, E. (1994) Chromosomal mapping and nucleotide sequence of two tandem repeats of Atlantic salmon 5S rDNA. *Cytogenet. Cell Genet.* **67**, 31–36.
- Pendas, A. M., Moran, P., Martinez, J. L., and Garcia-Vazquez, E. (1995) Applications of 5S rDNA in Atlantic salmon, brown trout, and in Atlantic salmon x brown trout hybrid identification. *Mol. Ecol.* **4**, 275–276.
- Quinn, J. S., Guglich, E., Seutin, G., Lau, R., Marsolais, J., Parna, L., Boag, P. T., and White, B. N. (1992) Characterization and assessment of an avian repetitive DNA sequence as an icterid phylogenetic marker. *Genome* **35**, 155–162.
- Sambrook, J. H., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., pp. 9.16–9.19. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Sedman, Iu. E., Shostak, N. G., Kuprianova, N. S., Serenkova, T. I., Fel'gengauer, P. E., Gimalov, F., Lind, A. I., Timofeeva, M. I., and Baev, A. A. (1989) Intragenomic polymorphism of the primary structure of 5S rRNA gene variants of the loach (*Misgurnus fossilis* L.). Determination of the transcriptional activity. *Mol. Biol. (Mosk)* **23**, 1295–1308. (in Russian).
- Shimizu, Y., Oshiro, T., and Sakaizumi, M. (1993) Electrophoretic studies of diploid, triploid, and tetraploid forms of the Japanese silver crucian carp, *Carassius auratus langsdorfii*. *Jpn. J. Ichthyol.* **40**, 65–75.
- Singer, M. (1982) Highly repeated sequences in mammalian genomes. *Int. Rev. Cytol.* **76**, 67–112.
- Suzuki, H., Moriwaki, K., and Sakurai, S. (1994) Sequences and evolutionary analysis of mouse 5S rDNAs. *Mol. Biol. Evol.* **11**, 704–710.
- Takasaki, N., Park, L., Kaeriyama, M., Gharrett, A. J., and Okada N. (1996) Characterization of species-specifically amplified SINEs in three *Salmonid* species-chum salmon, pink salmon, and kokanee: the local Environment of the genome may be important for the generation of a dominant source gene at a newly retroposed locus. *J. Mol. Evol.* **42**, 103–116.
- Weiner, A. M., Deininger, P. L., and Efstratiadis, A. (1986) Non-viral retroposons: genes, pseudogenes and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.* **55**, 631–661.
- Wijers, E. R., Zijlstra, C., and Lenstra, J. A. (1993) Rapid evolution of horse satellite DNA. *Genomics* **18**, 113–117.