J. Indian Inst. Sci., Jan.–Feb. 1989, **69**, 47–53.

A highly repeated DNA sequence in the fish *Cyprinus carpio*

R. K. MANDAL, PRASUN DUTTA* AND UTPAL DATTA
Genetic Engineering Unit, Department of Biochemistry, Bose Institute, Calcutta 700054, India.

Received on April 18, 1988.

Abstract

A tandemly repeated DNA sequence is visible as a ladder of bands on agarose gel electrophoresis of *Hind* III-digested *Cyprinus carpio* DNA. Several members of this repeated DNA family have been cloned and sequenced. The repeat unit has a monomeric length of 245 bp and comprises about 8% of the fish genome. Sequence analysis showed to be AT rich with some direct and some inverted-repeat nucleotide clusters.

Key words: *Cyprinus carpio*, *Hind* III family, tandem repeat, nucleotide sequence.

1. Introduction

Multiple copies of closely related DNA elements forming families of cross-hybridizing sequences are present in all eukaryotic organisms in proportions ranging from less than 1% to more than 66% of the genome[1,2]. These repetitive DNAs can be isolated as satellite bands by isopycnic centrifugation in CsCl alone or in some cases, in presence of base-specific ligands. Two types of highly repeated DNA families have been well characterized: the tandemly repeated satellite DNA, arranged as long stretches on specific chromosome loci, and the interspersed DNA comprised of long and/or short sequences throughout the genome as unlinked single copies[3,4]. Wide variations in sequences are found between highly repeated DNAs from different species indicating a rapid rate of evolution. In spite of wide occurrence of such DNAs, their function is still a matter of conjecture.

Highly repeated satellite DNAs have been extensively investigated in plants[5], invertebrates[6,7], amphibians[8] and mammals[9–11]. Although fishes comprise a large class vertebrates, evolutionarily more ancient than amphibians, no studies on organization and cloning of repetitive DNA in these animals have come to our notice. We report here for the first time the cloning and characterization of a member of highly repeated DNA from a fish, namely, the common carp *Cyprinus carpio*.

*Present address: Deptt of Chemistry and Biochemistry, Southern Illinois University at Carbondale, Carbondale, Illinois 62901, U.S.A.

2. Experimental

2.1. *Isolation of DNA*

Live specimens of the fish *Cyprinus carpio* var-communis (Bangkok strain) were obtained from the local market. DNA was isolated from saline-EDTA-washed erythrocytes by phenol extraction method as described. First, chromatin was isolated from saline-EDTA-washed erythrocytes by osmolysis with five volumes of distilled water followed by the addition of NaCl solution to a final concentration of 0.1 M and spinning down the chromatin. Pelleted chromatin was then suspended by homogenising in a buffer containing 50 mM Tris-HCl, pH 7.4, 150 mM NaCl and 25 mM EDTA in a loose-fitting Dounce homogenizer and lysed by adding SDS to 1% followed by 1 h proteinase k(50 μg/ul) digestion at 37°C. Viscous lysate was heated at 65°C for 5 min and extracted with equal volume of water-saturated phenol, phenol:chloroform (1:1, v/v) and chloroform. To the final aqueous layer, 2.5 volumes of ethanol was added, incubated at −20°C and centrifuged the precipitated DNA. The DNA was dried, dissolved in TE(10 mM Tris-HCl, pH 8.0, 1 mM EDTA) and reprecipitated with 0.3 M Na-acetate-2.5 volumes of ethanol. The DNA pellet was washed with 70% EtoH and dehydrated ethanol, air dried and dissolved in TE to make a final concentration of 0.5-1 mg DNA/ml which was determined by measuring absorbance at 260 nm.

2.2. *Restriction digestion and Southern blotting*

Five micrograms of total erythrocyte DNA was digested to completion with *Hind* III, *Eco* RI and *Bam* HI (Bethesda Research Laboratories, USA, under conditions prescribed by the supplier) and fractionated by electrophoresis through 0.8% agarose gel in 40 mM Tris-acetate, 2 mM EDTA, pH 8.0 and 1 μg/ml EtdBr, at 3 v/cm. The gels were photographed on a Fotodyne UV-300 transilluminator. The developed negatives were traced with an LKB-2202 scanning densitometer to quantitate the bands. The DNA separated on the gel was transferred on to nitrocellulose filter (Schleizer and Schultz) as described by Southern[12]. DNA blots were hybridized with ^{32}P-nick translated ^{13}245 bp *Hind* III repeat fragments (mixed) of *Cyprinus* DNA or with labelled insert of pCchr-3 clone in the presence of 50% formamide at 42°C for 16-20 h. Blots were washed once with 0.2 × SSC-0.1% SDS at 60°C and twice with the same washing solution at room temperature, dried and exposed to Agfa-Gevaert Curix RP1 X-ray films at −70°C.

2.3. *Cloning of the repetitive DNA*

Total fish DNA isolated from erythrocytes was digested to completion with *Hind* III and the resulting fragments were separated by electrophoresis on a 0.8% low-melting temperature agarose gel in Tris-acetate buffer. The band representing repetitive DNA of about 245 bp was eluted by phenol extraction[14] and ligated with *Hind* III-cut pBR 322 plasmid DNA. The recombinant DNA was used to transform *E. coli* HB101 using standard methods[14]. The transformants were screened by colony hybridization[15] using labelled 245 bp monomer DNA isolated from low-melting agarose as a probe. Plasmid DNA from

the positive clones (designated as pCchr for *Cyprinus carpio Hin*d III repeat) were isolated as described by Birnboim and Doly[16].

2.4. *DNA sequencing*

The nucleotide sequence of the pCchr-3 and pCchr-63 plasmid inserts were determined following Sanger's method[17] of dideoxy chain termination sequencing, after recloning of the repetitive DNA in the single-stranded M13-phage mp18. The colourless positive clones containing pCchr-3 and pCchr-63 inserts were selected from the plate containing the chromogenic substrate X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside; Sigma). For each plasmid two phages containing inserts of the complementary strands were isolated and used for sequencing. The ddNTPs and dNTPs were obtained from Sigma. The large fragment of *E. coli* DNA pol I (Klenow fragment) was from BRL and ^{35}S-dATP was from Amersham. The sequencing gel (8%) after electrophoresis was fixed in 10% MeOH-10% acetic acid solution for 30 min, drained out the fixing solution, transferred the gels to Whatman 3 mm paper, dried completely in vaccum gel drier (Atto) at 50°-60°C and exposed to Agfa-Gevaert Curix RP1 X-ray films at -70°C using intensifying screen for varying times.

3. Results and discussion

3.1. *Identification and cloning of the fish repetitive DNA*

*Hin*d III digests of cyprinus DNA showed a visible ladder of about 245 bp fragments on EtdBr staining after separation by 0.8% agarose gel electrophoresis (fig. 1A). On the other hand, *Eco*RI and *Bam*HI digests did not show such prominent bands (result not represented). On densitometric scanning of the negative of *Hin*d III gel pattern, the fastest moving monomer band amounted to about 2.8% of the total genomic DNA (fig. 1B). When the 245 bp monomer band isolated from low-melting point agarose was ^{32}P-labelled by nick translation and used as a probe in Southern blot hybridization of the *Hin*d III digest, a ladder of 245 bp repeats was revealed in the autoradiogram (fig. 1C), thus indicating a tandem repetition, the ladder being produced due to loss of the *Hin*d III site in some of the repeats.

The 245 bp *Hin*d III fragment of repetitive DNA was cloned in the plasmid pBR 322 and the clones were picked up by colony hybridization (fig. 2A). The recombinant plasmids were isolated from several positive clones and mapped with several restriction enzymes (fig. 2B, C). When a *Hin*d III digest of genomic DNA was probed with different clones identical ladder patterns appeared on Southern blotting indicating that the repeat units were very similar in size and sequence, except for an occasional base change (result not shown).

3.2. *Quantization of the Hin*d *III-repetitive DNA*

Different amounts of pCchr-3 and *Cyprinus* genomic DNA were separately digested with

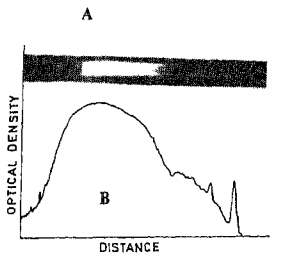

FIG. 1. Presence of 245 bp repeats in digests of *Cyprinus* DNA. (A) Five μg of blood DNA was digested to completion (5 units/μg DNA for 2 h) with *Hind* III, electrophoresed in 0.8% agarose gel and stained with EtdBr. (B) Densitometric scan of the stained gel as in A. (C) The *Hind* III-digested DNA separated in the gel as in panel A was Southern blotted on to nitrocellulose filter and hybridized with ^{32}P-labelled 245 bp *Hind* III repeat fragments (mixed) of *Cyprinus* DNA isolated from low-melting agarose.

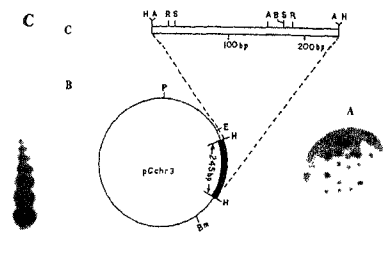

FIG. 2. Cloning of 245 bp *Hind* III repeat of *Cyprinus* DNA. The 245 bp *Hind* III-monomeric DNA was isolated from low-melting agarose, ligated to *Hind* III cut pBR322 and used to transform *E. coli* HB101. (A) Colony screening of the transformants. (B) The recombinant plasmid pCchr-3 containing the 245 bp repeat. (C) Restriction map of the 245 bp insert. The symbols are: A = *Alu*I; B = *Bcl*I; Bm = *Bam*HI; *E* = *Eco*RI; H = *Hind* III; P = *Pst*I; R = *Rsa*I; S = *Sau*3AI.

Hind III, separated on 0.8% agarose gel and blotted. The blot was hybridized with ^{32}P-labelled 245 bp insert of pCchr-3 and autoradiographed. By comparison of the densitometric scans (results not presented) the monomeric peak area given by 200 ng of genomic DNA was equal to that given by 6 ng of cloned-monomer insert of pCchr-3. Thus the amount of monomer is 3% of the DNA in the digest, a value comparable to that obtained by scanning of the EtdBr-stained gel (fig. 1B). Again, the integrated area under all the peaks in the autoradiogram of genomic DNA digest was 2.65 times than that under the monomeric peak. This leads to an amount of 8% as the repeated DNA in Cyprinus genome, corresponding to a copy number of 6×10^5, as cyprinus has a 'c' value of 1.7 pg[18].

3.3. *Nucleotide-sequence analysis*

The nucleotide sequences of two clones (pCchr-3 and pCchr-63) are presented in fig. 3. The repeat is 245 bp and contains several common 4-nucleotide restriction sites such as *Alu* I, *Dde* I, *Mbo* I, *RSA* I and *Sau* 3AI. The two sequences are identical except for a few nucleotide changes. The DNA is A + T rich (63%) with clusters of T and A. There are several direct repeats and inverted repeats as indicated in fig. 3, but none is more than ten nucleotides long. No internal subrepeats were found. In this respect the fish DNA repeat resembles the repetitive *Eco* RI monomer-3 sequence of *Xenopus* DNA[8] and *Hind* III repeat of *Arabidopsis* DNA[5]. An examination of open reading frames in the sequence following start codon ATG reveals frequent stop codons, suggesting that there might be no RNA translation even if DNA were transcribed. However, we failed to detect any RNA transcript corresponding to the repetitive DNA in Northern blots.

```
                                                          1          50
pCchr-3   AAGCTTTAGT CTTAACGTTT GTACAAACTA TCATTCTCTA ACAGAGAAAG
                              4        4'                      6

pCchr-63  ---------- ---------- ---------- ---------- ----------

                                          2                      100
pCchr-3   AAGGTTTTCA GCACTTTGTG GGCTTTCTTT CTGTTCATTT GCTTAGTTGC
          6'

pCchr-63  -----A---- ---------- ---------- -----A---- ----------

          1'        2'                    3                      150
pCchr-3   ACTAACAGAG TGTTTCTGTT CTCAGAAACG CTAAACTGAG CGTTTTTATG
                    5         5'

pCchr-63  ---------- --A-A----- -A-------- ---------- ----C-----

                                        3'                       200
pCchr-3   CTTAGAAGCT CAAACATGAG TTCATGATCA TAAACTAGTA CTCACTGAAC

pCchr-63  ------C--- G------AT- ---------- ----------

                                                                 250
pCchr-3   TGTTCTGCAT TGCATACATT CATTGAGATG TTAGACACTT ATTGCAAGCT
```

FIG 3 Nucleotide sequence of the monomeric insert from the recombinant plasmids pCchr-3 and pCchr-63 The sequences are presented from the 5′ to 3′ direction For the clone pCchr-63, only the nonhomologous nucleotides are shown, the homologous ones being indicated by hyphens. The three direct repeats (1, 1′; 2, 2′; 3, 3′) and three inverted repeats (4, 4′; 5, 5′; 6, 6′) are indicated by arrows over and under the sequences respectively. The restriction sites and open reading frames as determined by computer analysis are not indicated in the figure.

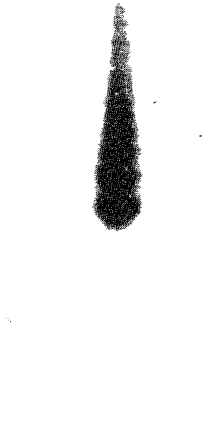

FIG. 4. The presence of repetitive DNA sequences homologous to *Cyprinus* 245 bp *Hind* III repeats in other fish species. Five μg of blood DNA from different fishes was separately digested with *Hind* III (20 units, 2 h), Southern blotted and probed with ^{32}P-labelled pCchr-3 DNA. (a) *Silver carp* (b) *Cyprinus carpio* (c) *Catla catla* (d) *Heteropneustes fossilis* (e) *Tilapia mossambica* (f) *Anabas testudineus.*

3.4. *Species and sex specificity of the repetitive sequences*

The representation of the 245 bp repeat-like DNA sequences in the genome of a number of fish was examined by hybridization of Southern blots of *Hind* III-digested genomic DNA with nick-translated pCchr-3. Under the hybridization condition similar but less intense ladder patterns were obtained with DNA from several fishes but not with that of others (fig. 4). This observation suggests that the fish repetitive DNA can be used as a marker to follow the evolutionary closeness among fish species.

As most fishes including cyprinus are sexually dimorphic but lack any differentiated sex chromosomes, we titrated the levels of the repetitive DNA in the genome of male and female cyprinus. There is no difference in copy number of the repeats in the two sexes (result not presented).

4. Conclusion

We report here the characterization of a highly repetitive DNA in the fish *Cyprinus carpio* which appeared as tandem repeats of 245 bp, each unit containing a *Hind* III site. The 6×10^5 copies of the repeat present per haploid genome comprised about 8% of the genome. The presence of direct and inverted repeat sequences was reminiscent of movable elements[20]. The fish sequences did not seem to be the result of amplification and divergence of a small consensus sequence, as are other common satellite DNAs in *Drosophila* or mouse[2]. This is the first report on the characterization of a highly repetitive family of nucleotide sequence in a member of Pisces, a large class of animals. The cloned repetitive DNA should be useful in evolutionary studies of related groups of fish.

Acknowledgements

We thank Dr B. B. Biswas of this Institute for his interest in this work, Dr Malay Ray of the Centre for Cellular and Molecular Biology, Hyderabad (India) and Dr. Ranjit Dasgupta of the University of Wisconsin (U.S.A.), for providing materials, strains and advice on nucleotide sequencing, and to the Department of Science and Technology, Government of India, for financial support.

References

1. Britten, R. J. and Kohne, D. E. Repeated sequences in DNA, *Science*, 1968, **161**, 529–540.
2. Brutlag, D. L Molecular arrangement and evolution of heterochromatic DNA, *Annu. Rev. Genet.*, 1980, **14**, 121–144.
3. Jelinek, W. R. and Schmid, C. W. Repetitive sequences in eukaryotic DNA and their expression, *Annu. Rev. Biochem.*, 1982, **51**, 813–844.
4. Singer, M F. Highly repeated sequences in mammalian genomes, *Int. Rev. Cytol.*, 1982, **76**, 67–112.

5. MARTINEZ-ZAPATER, J. M., ESTELLE, M. A. AND SOMERVILLE, C. R. A highly repeated DNA sequence in *Arabidopsis thaliana*, *Mol. Gen Genet.*, 1986, **204**, 417–423.

6. FOWLER, R. F. AND SKINNER, D. M. Cryptic satellites rich in inverted repeats comprise 30% of the genome of a hermit crab, *J. Biol. Chem.*, 1985, **260**, 1296–1303.

7. BARIGOZZI, C., BADARACCO, G., PLEVANI, P., BARATELLI, L., PROFETA, S., GINELLI, E. AND MENEVERI, R. Heterochromatin in the genus *Artemia*, *Chromosoma*, 1984, **90**, 332–337.

8. HUMMEL, S., MEYERHOF, W., KORGE, E. AND KNÖCHEL, W. Characterization of a highly and moderately repetitive 500 bp *Eco* RI fragments from *Xenoupus laevis* DNA, *Nucl. Acids Res.*, 1984, **12**, 4921–4938.

9. HÖRZ, W. AND ALTENBURGER, W. Nucleotide sequence of mouse satellite DNA, *Nucl. Acids Res.*, 1981, **9**, 683–696.

10. ARNASON, U., HOGLUND, M. AND WIDEGREN, B. Conservation of highly repetitive DNA in cetaceans, *Chromasoma*, 1984, **89**, 238–242.

11. NOVAK, U. Structure and properties of a highly repetitive DNA sequence in sheep, *Nucl. Acids Res.*, 1984, **12**, 2343–2350.

12. SOUTHERN, E. M. Detection of specific sequences among DNA fragments separated by gel electrophoresis, *J. Mol. Biol.*, 1975, **98**, 503–517.

13. RIGBY, P. W. J., DIECKMANN, M., RHODES, C. AND BERG, P. Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I, *J. Mol. Biol.*, 1977, **113**, 237–251.

14. MANIATIS, T., FRITSCH, E. F. AND SAMBROOK, J. *Molecular cloning–A laboratory manual*, Cold Spring Harbor Laboratory, 1982.

15. GERGEN, J. P., STERN, R. H. AND WENSINK, P. C. Filter replicas and permanent collections of recombinant DNA plasmids, *Nucl. Acids Res.*, 1979, **7**, 2115–2135.

16. BIRNBOIM, H. C. AND DOLY, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA, *Nucl. Acids Res.*, 1979, **7**, 1513–1523.

17. SANGER, F., NICKLEN, S. AND COULSON, A. R. DNA sequencing with chain-terminating inhibitors, *Proc. Natn. Acad. Sci. USA*, 1977, **74**, 5463–5467.

18. HINEGARDNER, R. AND ROSEN, D. E. Cellular DNA content and the evolution of teleostean fishes, *Am. Nat.*, 1972, **106**, 621–644.

19. THOMAS, P. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose, *Proc. Natn. Acad. Sci. USA*, 1980, **77**, 5201–5205.

20. KLECKNER, N. Transposable elements in prokaryotes, *Annu. Rev. Genet.*, 1981, **15**, 341–404.