Genetica (2006) 127:133–141
DOI 10.1007/s10709-005-2674-y

A tandemly repetitive centromeric DNA sequence of the fish *Hoplias malabaricus* (Characiformes: Erythrinidae) is derived from 5S rDNA

Cesar Martins[1], Irani Alves Ferreira[1], Claudio Oliveira[1], Fausto Foresti[1] & Pedro Manoel Galetti Jr[2]
[1]*Departamento de Morfologia, UNESP-Universidade Estadual Paulista, Instituto de Biociências, CEP 18618-000, Botucatu, SP, Brazil (E-mail: cmartins@ibb.unesp.br; Phone/Fax: +55-14-38116264);* [2]*Departamento de Genética e Evolução, Universidade Federal de São Carlos, Centro de Ciências Biológicas e da Saúde, CEP 13565-905, São Carlos, SP, Brazil*

Received 15 June 2005; Accepted 31 August 2005

Key words: 5S rDNA, 5S rDNA variant, centromeric DNA, fish, *Hoplias malabaricus*, satellite DNA

Abstract

A substantial fraction of the eukaryotic genome consists of repetitive DNA sequences that include satellites, minisatellites, microsatellites, and transposable elements. Although extensively studied for the past three decades, the molecular forces that generate, propagate and maintain repetitive DNAs in the genomes are still discussed. To further understand the dynamics and the mechanisms of evolution of repetitive DNAs in vertebrate genome, we searched for repetitive sequences in the genome of the fish species *Hoplias malabaricus*. A satellite sequence, named 5S*Hind*III-DNA, which has a conspicuous similarity with 5S rRNA genes and spacers was identified. FISH experiments showed that the 5S rRNA bona fide gene repeats were clustered in the interstitial position of two chromosome pairs of *H. malabaricus*, while the satellite 5S*Hind*III-DNA sequences were clustered in the centromeric position in nine chromosome pairs of the species. The presence of the 5S*Hind*III-DNA sequences in the centromeres of several chromosomes indicates that this satellite family probably escaped from the selective pressure that maintains the structure and organization of the 5S rDNA repeats and become disperse into the genome. Although it is not feasible to explain how this sequence has been maintained in the centromeric regions, it is possible to hypothesize that it may be involved in some structural or functional role of the centromere organization.

Introduction

An interesting feature of eukaryote genome is the presence of a substantial fraction of duplicated DNA sequences, most of them composed by non-coding sequences that include satellite, minisatellite and microsatellite sequences, and transposable elements. Although studied extensively for the past three decades, the molecular forces that propagate and maintain repetitive DNAs in the genome are still discussed. Among whole sequenced genomes the repetitive areas remains as gaps because of the difficulty in their correct positioning and array in the genome. However, the role of these DNAs in genome organization and evolution, and their impact on speciation has been frequently reported (Charlesworth, Snlegowski and Stephan, 1994). The variation in genome size of different eukaryotes is often reported to differences in the amount of repeated DNA sequences (Cavalier-Smith, 1985; Brenner et al., 1993). Recently advances on studies concerning non-coding repetitive DNA sequences have shown that such sequences are extremely important in the structural and functional organization of the genome.

Among vertebrate species, studies about repetitive sequences have been mainly directed to the understanding of the structure and organization of satellite DNA and ribosomal DNA (rDNA) repeats. Satellite DNA families can correspond to

10–20% of some mammalian genomes (Beridze, 1986) and are usually species-specific or found to be conserved within closely related species. They are useful for molecular cytogenetic analysis, such as the identification of homologous chromosomes and chromosomal abnormalities by *in situ* hybridization. The molecular organization, chromosomal location, and possible functions of satellite DNAs have been studied in several groups of animals (Singer, 1982; Clabby et al., 1996). These studies have indicated that satellite-like repetitive DNA sequences may play an important role at the chromosomal and nuclear level (Singer, 1982; Larin, Fricker and Tyler-Smith, 1994; Sart et al., 1997).

Although usually considered as a single biological species, the taxonomy of *Hoplias malabaricus* is poorly understood. Growing evidence has pointed to the karyotypic diversity of *H. malabaricus*, showing interpopulational differences in the diploid number and chromosome morphology, as well as in sex chromosome systems (Bertollo, Takahashi and Moreira-Filho, 1978, 1983; Dergam and Bertollo, 1990; Bertollo, Moreira-Filho and Fontes, 1997a, Bertollo et al., 1997b; Lopes et al., 1998; Bertollo and Mestriner, 1998; Born and Bertollo, 2000). Specimens with a putative hybrid karyotype have not been found when distinct chromosomal forms (cytotypes) are sympatric (Bertollo et al., 2000). Since the fish *H. malabaricus* has shown to be an interesting model species for cytogenetic and evolutionary studies, we have investigated repetitive sequences in the genome of this fish. A centromeric satellite DNA family originated from the 5S rDNA repeats was isolated and its composition and organization in the genome of *H. malabaricus* studied.

Material and methods

Animals, DNA isolation, DNA digestion, PCR and cloning

Genomic DNA of five individuals of *Hoplias malabaricus* from the Araquá river (Botucatu, SP, Brazil) was extracted according to standard phenol–chloroform procedures (Sambrook and Russel, 2001). Restriction enzyme digestions of the genomic DNA were conducted with the endonucleases *Hind*III, *Msp*I, *Pst*I, *Hae*III, *Eco*RI, *Pvu*II, *Sca*I, *Rsa*I and *Spe*I. The endonuclease *Hind*III generated a band of approximately 350 bp that was purified from agarose gel for cloning. PCR amplifications of 5S rDNA were performed using primers A (5′-TAC GCC CGA TCT CGT CCG ATC-3′) and B (5′-CAG GCT GGT ATG GCC GTA AGC-3′) designed from the 5S rRNA sequence of rainbow trout (Komiya and Takemura, 1979) to amplify the 5S rRNA gene and the non-transcribed spacer (NTS) (hereafter refereed as 5S rDNA repeat). A standard PCR reaction was performed using 150 pmoles of each primer, 20 ng of genomic DNA, 1x *Taq* buffer, 200 μM of each dNTP and 2 U of *Taq* polymerase in a final reaction volume of 50 μl. Cycling times were as follows: 94°C of denaturation for 5 min, 35 cycles of 95°C for 1 min, 63°C for 30 s and 72°C for 1 min, with post-cycling extension at 72°C for 7 min. The PCR-amplified products were visualized in 1% agarose gels. The *Hind*III-DNA fragments and the PCR-generated 5S rDNA repeats were linked in the plasmids pUC18 *Hind*III-BAP (Amersham Bioscience) and pGEM-T (Promega), respectively, and used to transform DH5α *E. coli* competent cells (Invitrogen).

Sequencing and sequence analysis

The clones obtained from the 5S rDNA–PCR products (five clones) and the *Hind*III-DNA fragments (five clones) were sequenced on an ABI Prism 377 DNA sequencer (Perkin-Elmer) with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer). Nucleotide sequences were subjected to BLASTN (Altschul et al., 1990) searches at the National Center for Biotechnology Information (NCBI), web site (http://www.ncbi.nlm.nih.gov/blast), and the sequence alignment was performed using Clustal W (Thompson, Higgins and Gibson, 1994) and checked by hands. UPGMA based phylogenetic analyses and Kimura's 2-parameter genetic distances were determined using the software MEGA 2.1 (Kumar et al., 2001). Bootstrap resembling (Felsenstein, 1985) was applied to assess support for individual nodes using 1000 replicates.

Southern blot hybridization

The genomic organization of the isolated sequences was investigated by Southern blot-hybridization.

Genomic DNA (8 μg) was partially (10 and 30 min) and overnight (14 h) digested with 45 U of the endonuclease *Hind*III, submitted to gel-electrophoresis in 1% agarose and Southern-transferred to Hybon-N nylon membrane (Southern, 1975). The enzyme *Hind*III was used since has only one cleaving site in both isolated sequence classes. Filters containing the immobilized DNA were probed with cloned monomeric units of the PCR-generated 5S rDNA and with *Hind*III-DNA sequences. To avoid cross-hybridization between the 5S rDNA and the *Hind*III-DNA, the filter hybridization was carried out with the kit ECL-direct nucleic acid labeling and detection system (Amersham Bioscience) using high stringency conditions. The probes were denatured at 95°C, covalently labeled with the enzyme horseradish peroxidase and hybridized for 14 h to filter immobilized target DNA in a hybridization buffer (6 M urea/50% formamide/0.5 M NaCl) at 42°C. After the hybridization the filters were washed in 6 M urea/0.4% SDS/0.5×SSC buffer at 42°C and the hybridized DNA detected by chemiluminescence.

Chromosome analyses

Mitotic chromosomes were prepared from anterior kidney cells with *in vivo* colchicine treatment (Bertollo, Takahashi and Moreira-Filho, 1978) and were submitted to fluorescence *in situ* hybridization (FISH) (Pinkel et al., 1986) and C-banding (Sumner, 1972). The same probes employed in the Southern hybridizations (5S rDNA repeat and *Hind*III-DNA fragment) were used for FISH. The probes were labeled by nick translation with biotin-14-dATP (Bionick labelling system-Invitrogen). The hybridization was performed in three stringency conditions: 35, 50 or 65% of formamide. The metaphase chromosome slides were incubated with RNase (40 μg/ml) for 1.5 h at 37°C. After denaturation of chromosomal DNA in 70% formamide/2×SSC for 4 min at 70°C, hybridization mixtures containing 100 ng of denatured probe, 10 mg/ml dextran sulfate, 2×SSC and 35, 50 or 65% of formamide were dropped on the slides and the hybridization was performed overnight at 37°C. Hybridization washes included 2×SSC and 35, 50 or 65% of formamide at 37°C and 2×SSC and 4×SSC at room temperature. Detection of hybridized probes was carried out with avidin–FITC conjugate (Sigma) followed by two rounds of signal-amplification. After each amplification step, the slides were washed in the blocking buffer (1.26% $NaHCO_3$, 0.018% sodium citrate, 0,0386% triton, 1% non-fat dried milk) at 42°C. Chromosomes were counterstained with propidium iodide, and the slides were mounted with antifade (Oncor).

Results

The cloning and sequencing of the repetitive 5S*Hind*III-350 bp band identified DNA fragments of 356–360 bp long with the presence of insertions, deletions and base substitutions among the clones (Figure 1). A search in the DDBJ-EMBL-GenBank database showed DNA sequence similarity with the 5S rRNA gene of several vertebrates, including fishes. To determine the relationship between the repetitive *Hind*III-DNA sequence and the 5S rRNA gene, PCR with specific primers designed to amplify monomeric repeat units of the 5S rDNA from the genome of *H. malabaricus*, was performed. The primers designed successfully amplified the incomplete 5S rRNA gene (118 bp) and the complete NTS (Figure 1). The 5S rDNA consensus sequence and the *Hind*III-repetitive sequences (hereafter denominated *Hind*III-DNA) were aligned and the similarity between them determined (Figure 1). Analyses among the 5S*Hind*III-DNA sequences showed a mean genetic distance of 0.049. The genetic distance between the 5S rDNA and the 5S*Hind*III-DNA sequences ranges from 0.173 (most similar sequences) to 0.221 (most divergent sequences) (Table 1). The phylogenetic analysis of the sequences discriminated the 5S rDNA and the 5S*Hind*III-DNA sequences (Figure 2). The AT base content of 5S rDNA repeat units was 55.9% and the 5S*Hind*III-DNA 65.48%. The main difference between 5S rDNA repeat units and the 5S*Hind*III-DNA is the presence of an expanded imperfect TAAA microsatellite sequence and two short deletions.

A GenBank search showed that the 5S rDNA and the 5S*Hind*III-DNA sequences isolated in the present work had similarities with a satellite DNA family previously identified in the genome of *H. malabaricus* (Database accession numbers L11927 and L11928) (Haaf et al., 1993).

```
5SrDNA       TTAGCTTAAG C-TCC----- --TAAAGACA -GTAGCCACA GTCATTTAAA CACAAGAGAA GAAGAAAAC-
5SHindIII-a ...TT..TC. GC...GATTG --.G....-T T......... .......... .......... ..GC.....T
5SHindIII-b .C...AA... GC...GATTG --.G....-T T...C.T... .......... .......... ..GC.....T
5SHindIII-c .......T.. GC...GATTG --.G..A.-T T...C.T... .......... .......... ..GC...C.T
5SHindIII-d ..C..AA... GC...GATTG -T.G..A.TT T...C.T... .......G.. .......... ..GC.....T
5SHindIII-e .CC..AA... GC...GATTT GTCG....TT T...C.T... ..T....G.. .......C.. A.GC.....T

5SrDNA       ---------- ---------- ---------- ---------- ---------- TTAAAAAAAA AAA-AAACGA
5SHindIII-a TTAAAAAATA AGCAAATAAA CAATGAATAA ATAAATAATA AACAAACAAA .A.......C ...T....A.
5SHindIII-b TTAAAAAATA AGCAAATAAA CAATGAATAA ATAAATAATA AACAAACAAA .A.......C ...T....A.
5SHindIII-c TTGAAAAATA AACAAATAAA CAATAAATAA ATAAAAAATA AACAAACAAA .A.......C ...T....A.
5SHindIII-d TTGAAAAATA AACAAATAAA CAATAAATAA ATAAAAAATA AACAAACAAA .A.G.....C ...T......
5SHindIII-e TTAAAAAATA AACAAATAAA CAATGAATAA ATAAAGAATA AACAAACAAA .A.......C ...T......

5SrDNA       A-----CTGG AAAGCCAATC ATTTTCATTT GAGGTGGCAG GCAGTCTACC GTTTACAGGC AGCTCTAAAA
5SHindIII-a .AAACA.... .......... .....A.... ...T..TG.. ..TT.T..-- .......A.. ......GT..
5SHindIII-b .AAACA.... .......... .....A.... ...T..TG.. ..TT.T..-- .......A.. ......GT..
5SHindIII-c .AAACA.... .......... .....A.... ...T..TG.. ..TT.T..-- .......A.. ......GT..
5SHindIII-d .AAACA.... .......... .....A.... ...T..TG.. ..TT.T..-- .......A.. ......GT..
5SHindIII-e .AAACA.... .......... ..G..A.... ...T..TG.. ..TT.TC.-- .......A.. ......GT..
                                                                                ←5SB
5SrDNA       CAAGTATTAG TCTTAGAAGA GTCTCTAGAA ATAGTCTTCT ACAGCCAGCA GTGATCGCTT ACGGCCATAC
5SHindIII-a ....AG...C ........C. ...AT..C.. .......... .......A.. ..C..T.... .T.....G..
5SHindIII-b ....AG...C ........C. ...AT..C.. .......... .......A.. ..C..T.... .T.....G..
5SHindIII-c ....AG...C ........C. ...AT..C.. .......... .......A.. ..C..T.... .T........
5SHindIII-d ....AG...C ........C. ...AT..C.. .......... .......A.. ..C..G.... .T........
5SHindIII-e ....AG...C ........C. ...AT..C.. .......G.. .......A.. ..C..T.... .T........
                                  5SA→
5SrDNA       CAGCCTGNNT ACGCCCGATC TCGTCCGATC TCGGAAGCGA AGCAGAGTTG GGCCTAGTTA GTACTTGGAT
5SHindIII-a .CCT..---. .-.AA--... ----...... .T.------- --------.. C..T...... .........A
5SHindIII-b .CCT..---. .-.AA--... ----...... .T.------- --------.. C..T...... .........A
5SHindIII-c .CCT..---. .-.AA--... ----...... .T.------- --------.. C..T...... .........A
5SHindIII-d .CCT..---. .-.AA--... ----...... .T.------- --------.. C..T...... .........A
5SHindIII-e .CCT..---. .-.AA--..G ----...... .T.------- --------.. C..TG..... .........A

5SrDNA       GGGA-GACTG CCTGGGAATA CTAGGTGCTG TAAGCTT-318
5SHindIII-a T...-..... ..A....... .......... .......-356
5SHindIII-b T...-..... ..A....... .......... .......-356
5SHindIII-c T.A.-..... ..A.....CT .......... .......-356
5SHindIII-d T.A.-..... ..A.....C. .......... .......-358
5SHindIII-e T...A..... ..A....... .......... .......-360
```

Figure 1. Alignment of the 5S rDNA consensus sequence and 5S*Hind*III-DNA repeat units (5S*Hind*III-a to 5S*Hind*III-e) identified in the genome of *H. malabaricus*. The 5S rRNA gene sequence regions are in boldface type, the primers used to obtain the 5S rDNA sequences are underlined and the arrows indicate the direction of the PCR amplification. Hyphens represent gaps, dots base identity, and N non-defined nucleotides. The sequences are deposited in GenBank under the accession numbers AY624052-AY624061.

However, no previous relationship was described between that satellite sequence and the 5S rDNA sequences.

Southern blot hybridization analyses were conducted using DNA digested with the restriction endonuclease *Hind*III, selected for its recognition site within the repetitive isolated sequences. The enzyme *Hind*III has only one cleaving site in both isolated sequences. Membrane hybridization using the 5S*Hind*III-DNA and 5S rDNA isolated sequences as probes showed that these repetitive sequences are tandemly arrayed in the genome of *H. malabaricus* (Figure 3).

FISH was carried out using the 5S rDNA and the 5S*Hind*III-DNA sequences as probes. Under low stringent conditions (35% of formamide) both probes hybridized in the centromeric region of 18 chromosomes and near the centromeres in the short arm of chromosome pairs 3 and 15 (Figure 4). Under high stringent conditions (50 and 65% formamide) the 5S*Hind*III-DNA probe hybridized to the centromeric region of 18 chromosomes and

Table 1. Genetic distances (Kimura's 2-parameter) determined between 5S rDNA consensus sequence and 5S*Hind*III-DNA sequences (*Hind*III-a to *Hind*III-e). Upper diagonal, standard error; lower diagonal, genetic distances

	5S rDNA	5S*Hind*III-a	5S*Hind*III-b	5S*Hind*III-c	5S*Hind*III-d	5S*Hind*III-e
5S rDNA	–	0,025	0,025	0,027	0,027	0,029
5S*Hind*III-a	0,173	–	0,008	0,011	0,012	0,014
5S*Hind*III-b	0,177	0,026	–	0,010	0,010	0,112
5S*Hind*III-c	0,186	0,043	0,40	–	0,009	0,015
5S*Hind*III-d	0,194	0,62	0,41	0,029	–	0,013
5S*Hind*III-e	0,221	0,74	0,46	0,074	0,055	–

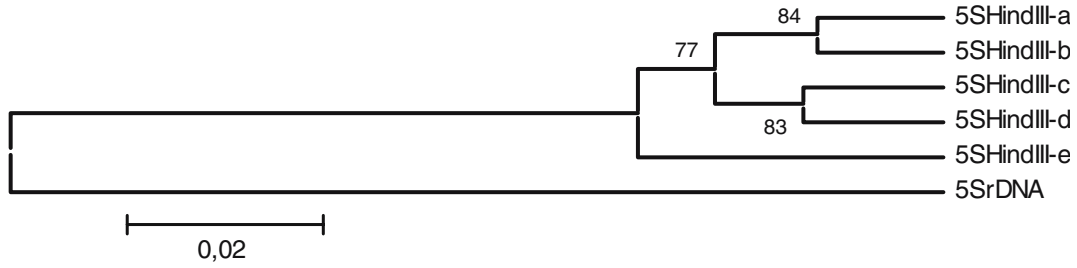

Figure 2. UPGMA three based on the 5S rDNA consensus and 5S*Hind*III-DNA sequences (5S*Hind*III-a to 5S*Hind*III-e). The numbers at each node indicate the percentage recovery (> 50%) of the particular node (1000 bootstrap replicates).

the 5S rDNA probe to the short arm of chromosome pairs 3 and 15 (Figure 4). Dark C-band positive segments were detected in the centromeric region of most chromosomes whereas faint C-band positive segments were evidenced in the terminal position of several chromosomes as well as in the centromeres of a few chromosomes (Figure 4).

Discussion

Ohno, Wolf and Atkin (1968) postulated that gene duplication was the main driving force of vertebrates evolution. Once a gene was duplicated, one copy was no longer constrained by selection and any mutation that occurred in the duplicated copy could potentially lead to new expression patterns or altered function, leaving the original copy to provide its required function. The genetic studies, including the human genome sequencing, over the last years have identified that duplications of the genome have led to the complexity of human genes when compared to flies and worms (Horvath et al., 2001). In the present paper a tandemly repetitive centromeric DNA sequence (named 5S*Hind*III-DNA) from the fish *H. malabaricus* that share sequence similarities with repeat units of the 5S rDNA is reported. It seems probably that duplicated segments of the 5S rDNA gave origin to the 5S*Hind*III-DNA satellite family.

The presence of variant repeats of the 5S*Hind*III-DNA suggests that these sequences have an intense evolutionary rate in the genome. An evidence of the intense dynamism of the 5S*Hind*III-DNA sequences is the presence of the expanded TAAA microsatellite. The organization and evolution of tandem repetitive DNAs is governed by particular patterns of evolution such as unequal exchange, transposition, RNA-mediated transposition and gene conversion (Dover, 1986). Drouin and Moniz de Sá (1995) suggested the hypothesis that RNA-mediated transposition is the mechanism responsible for the unusual linkage of 5S rRNA genes to other tandemly repeated multigene families. According to the authors, the RNA-mediated transposition could be responsible for the dispersion of single copies of 5S rDNA repeats whereas covalently closed circular DNA (cccDNA) molecules containing 5S rRNA genes would be expected to sometimes lead to the insertion of several 5S RNA gene copies within others sequences in the genome. Such cccDNA

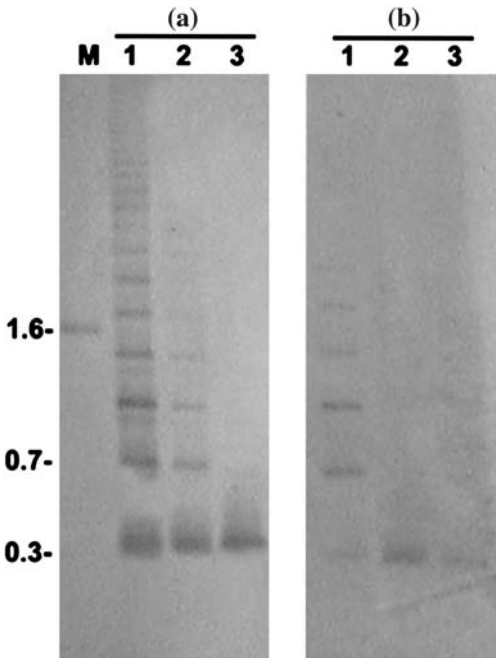

Figure 3. Southern blot hybridization of 5S*Hind*III-DNA (a) and 5S rDNA (b) sequences to genomic DNA of *H. malabaricus* digested with the endonucleases *Hind*III. 1 and 2 are partially digested genomic DNA samples (10 and 30 min, respectively), and 3, a totally digested DNA sample. M, molecular weight marker (kb).

molecules have been found in many eukaryotes species, including mammals, chicken, *Drosophila*, and plants (reviewed in Renault, DeGroote and Picard, 1993). Several classes of cccDNAs have been found in *D. melanogaster* embryos, one of them contains a variable number of sequences homologous to 5S rRNA genes (Pont, DeGroote and Picard, 1987). The first copies of the 5S*Hind*III-DNA of *H. malabaricus* could have transferred to centromeric position throughout any of the mechanisms described above or, alternatively, the 5S*Hind*III-DNA satellite sequence could have originated in the centromeric region of chromosomes 3 or 15 by duplication or chromosome inversion that included some adjacent 5S rDNA copies present in these chromosomes. The first 5S*Hind*III-DNA copies might have been associated with other repetitive sequences in the centromeric heterochromatin that facilitated its dispersion to the other chromosomes due to concerted evolution mechanisms.

Previous studies have found evidence of dispersed or clustered elements similar to 45S rDNA in several eukaryotic genomes. These elements have mainly been characterized as noncoding, small-unit tandem repeats of variable copy number. Such elements have been identified in various eukaryotic species, including yeast (Childs et al., 1981), animals (Lohe and Roberts, 1990), and plants (Falquet, Creusot and Dron, 1997). The results presented here for the fish *H. malabaricus* are showing that similar elements may be originated from the 5S rDNA. Dispersed 5S rDNA variants and pseudogenes seem to be common in mammals (Emerson and Roeder, 1984; Leah et al., 1990). On the other hand, the interesting feature of the variant 5S rDNA repeats of *H. malabaricus* is the high abundance of copy number, the tandem array, and their centromeric positioning.

If the 5S*Hind*III-DNA repeats may be conferring some structural or functional advantage to the chromosomes as a component of the centromeric DNA in *H. malabaricus* it cannot be addressed yet. Centromeres have been recognized as evolutionary dynamic regions of the genome (Eichler and Sankoff, 2003) but, although they have been well investigated from animals to fungus, important points remain to be understood (Henikoff, Ahmad and Malik, 2001). The centromere is vital for the correct sorting of chromosomes during cell division, being essential for the appropriated maintenance and segregation of the genetic material. Although this role is conserved throughout evolution, the DNA sequences found at centromeric regions are often variable (Henikoff, Ahmad and Malik, 2001). Disturbances in the structural and functional organization of the centromeres are critical leading to problems such as development defects and cancer. The centromeric regions are rich in repetitive DNAs, which is a trait of the characterized centromere regions from humans, mice, maize, fruit fly, and yeast (Henikoff, Ahmad and Malik, 2001). An extensive analysis of centromeric satellite DNAs of vertebrates showed the presence of short A-rich motifs, typical characteristics of centromere satellite (Vinãs et al., 2004). The expanded TAAA motif in the 5S*Hind*III-DNA is similar to the short A-rich motifs that were previously identified in the centromeric satellite DNAs of different fish species (Wright, 1989; Denovan and Wright, 1990; Garrido-Ramos et al., 1995; Kato, 1999; Canapa et al., 2002; Viñas et al., 2004). These short sequences are quite similar showing considerable similarity to other

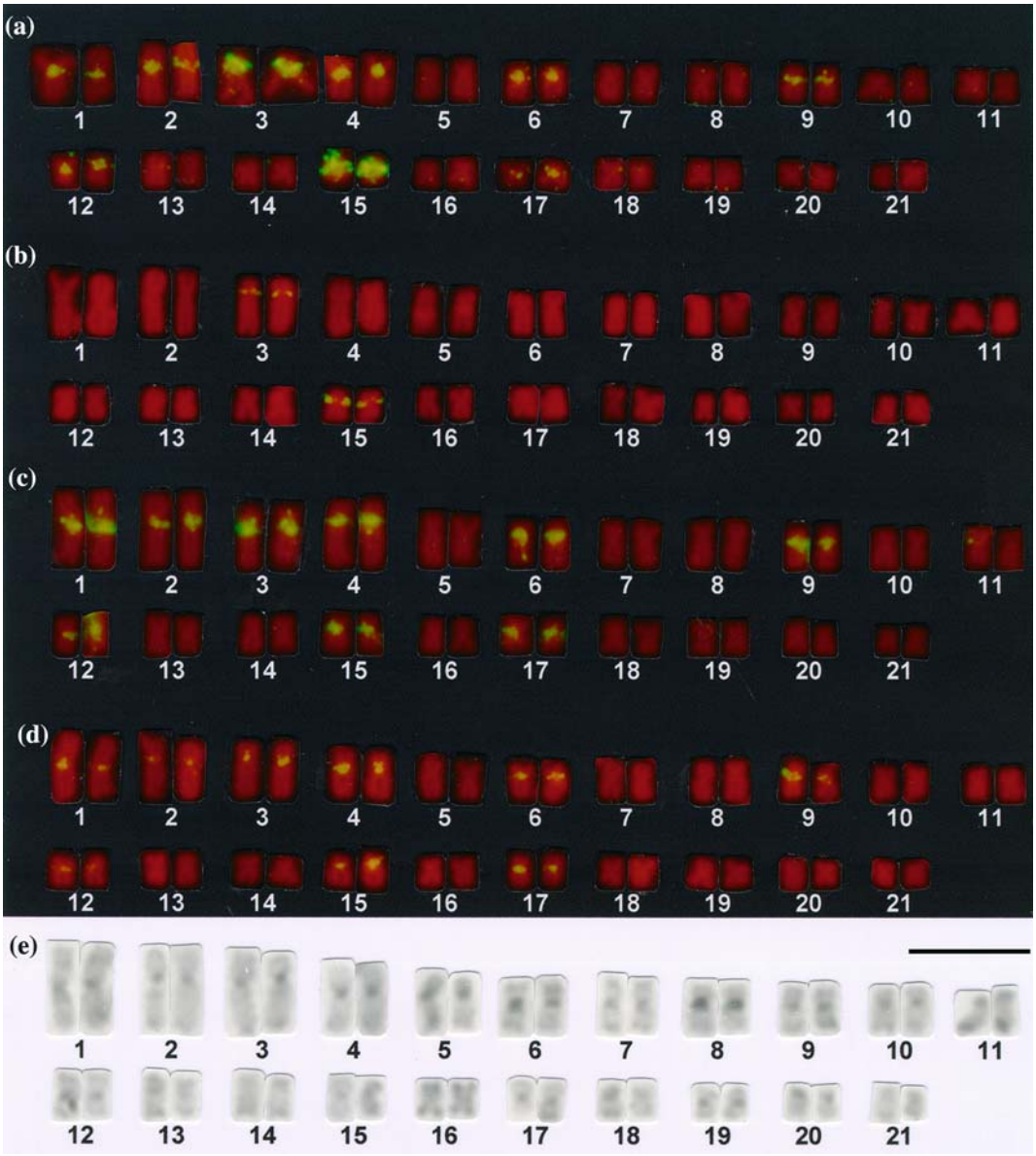

Figure 4. Karyotypes of *H. malabaricus* probed with 5S rDNA (a and b) and 5S*Hind*III-DNA (c and d) sequences (yellow signals) under lower stringent conditions (a and c) and higher stringent conditions (b and d). The heterochromatin distribution was visualized after treatment of the chromosomes with C banding (e). Bar = 5 μm.

centromeric motifs found in human (Vissel, Nagy and Choo, 1992), mouse (Wong and Rattner, 1988), and reptiles (Cremisi et al., 1988), suggesting that such sequences might play any important role in the structure and function of the *H. malabaricus* centromere.

Repetitive DNA sequences are subject to the action of several molecular mechanisms and are thought to be the most rapidly evolving components of eukaryotic genomes. The results described here also represent a good example of the fluidity of repetitive sequences providing novelties on the genomic organization of the centromeric region of vertebrates. Studies on the repetitive sequences can provide interesting insights for the comprehension of the genome structuring and evolution. The satellite 5S*Hind*III-DNA family has spread out in the centromeric region of several chromosomes

and has been favored during the evolution due to a possible role of the repetitive sequences of the centromeric region for the centromere structure and function.

Acknowledgements

The authors thank FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for financial support.

References

Altschul, S.F., W. Gish, W. Miller, E.W. Myers & D.J. Lipman, 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403–410.

Beridze, R., 1986. Satellite DNA. Springer, Berlin, Heidelberg, New York.

Bertollo, L.A.C. & C.A. Mestriner, 1998. The X1X2Y sex chromosome system in the fish *Hoplias malabaricus*. II. Meiotic analyses. Chromosome Res. 6: 141–147.

Bertollo, L.A.C., C.S. Takahashi & O. Moreira-Filho, 1978. Karyotypic studies of two allopatric populations of the genus *Hoplias* (Pisces, Erythrinidae). Brazil. J. Genet. 2: 17–37.

Bertollo, L.A.C., C.S. Takahashi & O. Moreira-Filho, 1983. Multiple sex chromosomes in the genus *Hoplias* (Pisces, Erythrinidae). Cytologia 48: 1–12.

Bertollo, L.A.C., O. Moreira-Filho & M.S. Fontes, 1997a. Karyotypic diversity and distribution in *Hoplias malabaricus* (Pisces, Erythrinidae): Cytotypes with 2n. 40 chromosomes. Brazil. J. Genet. 20: 237–242.

Bertollo, L.A.C., M.S. Fontes, A.S. Fenocchio & J. Cano, 1997b. The X1X2Y sex chromosome system in the fish *Hoplias malabaricus*. I. G-, C- and chromosome replication banding. Chromosome Res. 5: 493–499.

Bertollo, L.A.C., G.G. Born, J.A. Dergam, A.S. Fenocchio & O. Moreira-Filho, 2000. A biodiversity approach in the Neotropical fish *Hoplias malabaricus*. Karyotypic survey, geographic distribution of cytotypes and cytotaxonomic considerations. Chrom. Res. 8: 603–613.

Born, G.G. & L.A.C. Bertollo, 2000. An XX/XY sex chromosome system in a fish species, *Hoplias malabaricus*, with a polymorphic NOR-bearing X chromosome. Chromosome Res. 8: 111–118.

Brenner, S., G. Elgar, R. Sandford, A. Macrae, B. Venkatesh & S. Aparicio, 1993. Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. Nature 6452: 265–268.

Canapa, A., P.N. Cerioni, M. Barucca, E. Olmo & V. Caputo, 2002. A centromeric satellite DNA may be involved in heterochromatin compactness in gobiid fishes. Chrom. Res. 10: 297–304.

Cavalier-Smith, T., 1985. The Evolution of Genome Size. Wiley, New York.

Charlesworth, B., P. Snlegowski & W. Stephan, 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371: 215–220.

Childs, G., R. Maxson, R.H. Cohn & L. Kedes, 1981. Orphons: dispersed genetic elements from tandem repetitive genes of eukaryotes. Cell 23: 651–663.

Clabby, C., U. Goswami, F. Flavin, N.P. Wilkins, J.A. Houghton & R. Powell, 1996. Cloning, characterization and chromosomal location of a satellite DNA from the Pacific oyster, *Crassostrea gigas*. Gene 168: 205–209.

Cremisi, F., R. Vignali, R. Batistoni & G. Barsacchi, 1988. Heterochromatic DNA in *Triturus* (Anphibia, Urodela) II. A centromeric satellite DNA. Chromosoma 97: 204–211.

Denovam, E.M. & J.M. Wright, 1990. A satellite DNA family from pollock (*Pollachius virens*). Gene 87: 279–283.

Dergam, J.A. & L.A.C. Bertollo, 1990. Karyotypic diversi¢cation in *Hoplias malabaricus* (Osteichthyes, Erythrinidae) of the São Francisco and Alto Paraná basins, Brazil. Brazil. J. Genet. 13: 755–766.

Dover, G.A., 1986. Linkage disequilibrium and molecular drive in the rDNA gene family. Genetics 122: 249–252.

Drouin, G. & M. de Moniz Sá, 1995. The concerted evolution of 5S ribosomal genes linked to the repeat units of other multigene families. Mol. Biol. Evol. 12: 481–493.

Eichler, E.E. & D. Sankoff, 2003. Structural Dynamics of eukaryotes chromosome evolution. Science 301: 793–797.

Emerson, B.M. & R.G. Roeder, 1984. Isolation and genomic arrangement of active and inactive forms of mammalian 5S RNA genes. J. Biol. Chem. 259: 7916–7925.

Falquet, J., R. Creusot & M. Dron, 1997. Molecular analysis of DNA homologous to IGS subrepeats. Plant Physiol. Biochem. 35: 611–622.

Felsenstein, J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783–791.

Garrido-Ramos, M.A., M. Jamilena, R. Lozano, C. Ruiz Rejón & M. Ruiz Rejón, 1995. The *Eco*RI centromeric satellite DNA of the Sparidae family (Pisces, Perciformes) contains a sequence motive common to other vertebrates centromeric satellite DNAs. Cytogen. Cell Genet. 71: 345–351.

Haaf, T., M. Schmid, C. Steinlein, P.M. Galetti & H. Willard, 1993. Organization and molecular cytogenetics of a satellite DNA family from *Hoplias malabaricus* (Pisces, Erythrinidae). Chromosome Res. 1: 77–86.

Henikoff, S., K. Ahmad & H.S. Malik, 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293: 1098–1102.

Horvath, J.E., J.A. Bailey, D.P. Locke & E.E. Eichler, 2001. Lessons from the human genome: transitions between euchromatin and heterochromatin. Hum. Mol. Genet. 10: 2215–2223.

Kato, M., 1999. Structural bistability of repetitive DNA elements featuring CA/TG dinucleotide steps and mode of evolution of satellite DNA. Eur. J. Biochem. 265: 204–209.

Komiya, H. & S. Takemura, 1979. Nucleotide sequence of 5S ribosomal RNA from rainbow trout (*Salmo gairdnerii*) liver. J. Biochem. 86: 1067–1080.

Kumar, S., K. Tamura, I. Jakobsen & M. Nei, 2001. MEGA: Molecular evolutionary genetic analysis. Bioinformatics 12: 1244–1249.

Larin, Z., M.D. Fricker & C. Tyler-Smith, 1994. *De novo* formation of several features of a centromere following

introduction of a Y alphoid YAC into mammalian cells. Hum. Mol. Genet. 3: 689–695.

Leah, R., S. Frederiksen, J. Engberg & P.D. Sorensen, 1990. Nucleotide sequence of a mouse 5S rRNA variant gene. Nucl. Acids Res. 18: 7441–7441.

Lohe, A.R. & P.A. Roberts, 1990. An unusual Y chromosome of *Drosophila simulans* carrying amplified rDNA spacer without RNA genes. Genetics 125: 399–406.

Lopes, P.A., A.J. Alberdi, J.A. Dergam & A.S. Fenocchio, 1998. Cytotaxonomy of *Hoplias malabaricus* (Osteichthyes, Erythrinidae) in the Aguapey river (Province of Corrientes, Argentina). Copeia 1998: 485–487.

Ohno, S., U. Wolf & N.B. Atkin, 1968. Evolution from fish to mammals by gene duplication. Hereditas 59: 169–187.

Pinkel, D., T. Straume & J. Gray, 1986. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc. Natl. Acad. Sci. 83: 2934–2938.

Pont, G., F. DeGroote & G. Picard, 1987. Some extrachromosomal circular DNAs from *Drosophila* embryos are homologous to tandemly repeated genes. J. Mol. Biol. 195: 447–451.

Renault, S., F. DeGroote & G. Picard, 1993. Identification of short tandemly repeated sequences in extrachromosomal circular DNAs from *Drosophila melanogaster* embryos. Genome 36: 244–254.

Sambrook, J. & D.W. Russel, 2001. Molecular Cloning. A Laboratory Manual Third Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sart, D., M.R. Cancilla, E. Earle, J.I. Mao, R. Saffery, K.M. Tainton, P. Kalitsis, J. Martyn, A.E. Barry & K.H. Choo, 1997. A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. Nature Genet. 16: 144–153.

Singer, M.F., 1982. Highly repetitive sequences in mammalian genomes. Int. Rev. Cytol. 76: 67–112.

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

Sumner, A.T., 1972. A simple technique for demonstrating centromeric heterochromatin. Exp. Cell Res. 75: 304–306.

Thompson, J.D., D.G. Higgins & T.J. Gibson, 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22: 4673–4680.

Viñas, A., M. Abuín, B.G. Pardo, P. Martínez & L. Sánchez, 2004. Characterization of a new *Hpa*I centromeric satellite DNA in *Salmo salar*. Genetica 121: 81–87.

Vissel, B., A. Nagy & K.H.A. Choo, 1992. A satellite III sequence shared by human chromosomes 13, 14 and 21 that is contiguous with alpha satellite DNA. Cytogen. Cell Genet. 61: 81–86.

Wong, A.K.C. & J.B. Rattner, 1988. Sequence organization and cytological localization of the minor satellite of mouse. Nucl. Acids Res. 16: 11645–11661.

Wright, J.M., 1989. Nucleotide sequence, genomic organization and the evolution of a major repetitive DNA family in tilapia (*Oreochromis mossambicus*/*hornorum*). Nucl. Acids Res. 17: 5071–5079.