Original Article

Cytogenetic and Genome Research

Cytogenet Genome Res 2009;125:224–234
DOI: 10.1159/000230006

Accepted after revision: May 13, 2009
by M. Schmid

Organization of Repeated DNA Elements in the Genome of the Cichlid Fish *Cichla kelberi* and Its Contributions to the Knowledge of Fish Genomes

W.G. Teixeira[a] I.A. Ferreira[a] D.C. Cabral-de-Mello[a] J. Mazzuchelli[a]
G.T. Valente[a] D. Pinhal[a] A.B. Poletto[a] P.C. Venere[b] C. Martins[a]

[a]Departamento de Morfologia, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, SP,
[b]Instituto Universitário do Araguaia, Universidade Federal de Mato Grosso, Pontal do Araguaia, MT, Brazil

Key Words
Chromosome evolution · Cichlidae · Cytogenetics · Evolution · Genome · Heterochromatin · Transposons

Abstract
Repeated DNA elements have been extensively applied as physical chromosome markers in comparative studies for the identification of chromosomal rearrangements, the identification of sex chromosomes, chromosome evolution analysis and applied genetics. Here, we report the characterization of the transposable elements (TE) *Tc1*, *Rex1*, *Rex3* and *Rex6* and a new element called *RCk* in the genome of the South American cichlid fish *Cichla kelberi* using nucleotide sequence analysis and hybridization to metaphase chromosomes. The analysis of the repeated elements demonstrated that they are, in most cases, compartmentalized in heterochromatic regions, as has been observed in several other vertebrates. On the other hand, the elements *Rex1* and *Rex3* were also observed spanning extensive euchromatic regions on 2 chromosome pairs. The *RCk* element exhibits a wide distribution among fishes and also in amphibians, and it was spread throughout the chromosomes of *C. kelberi*. Our results have demonstrated that the compartmentalization of repeated elements is not restricted to heterochromatic segments, which has provided new concepts with regard to the genomic organization of transposons.

In most genomes, repeated sequences comprise a large portion of the DNA content of the cells, and the variation in the genome size of different eukaryotes is often attributed to differences in the amount of these sequences [Cavalier-Smith, 1985; Brenner et al., 1993]. Recent advances in the study of non-coding, repeated DNA elements have shown that such sequences are extremely important to the structural and functional organization of the genome [Grady et al., 1992; Schueler et al., 2001], and that they play important and specific roles in various cellular processes [Wong and Choo, 2004]. Among the completely sequenced genomes, the repeated elements remain as gaps, due to the difficulty of correctly identifying their position and array within the genome. Even the chromosomes that have been reportedly 'sequenced to completion', have multiple gaps in their pericentromeric regions related to the presence of duplicated and repeated segments [Horvath et al., 2001]. A better understanding of the genome structure and function requires the comprehension of these repeated segments. In addition, the integration of DNA sequences with physical chromosome mapping of these repeated elements can provide a better landscape of the genome, which is not yet clearly defined even in those genomes that are completely sequenced. The repeated DNA sequences can also serve as good chromosome markers, which could be useful in studies of species evolution,

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1424–8581/09/1253–0224$26.00/0

Accessible online at:
www.karger.com/cgr

Cesar Martins
Departamento de Morfologia, Instituto de Biociências
UNESP – Universidade Estadual Paulista, Botucatu, SP 18618-000 (Brazil)
Tel./Fax +55 14 3811 6264, E-Mail cmartins@ibb.unesp.br

the identification of chromosomal rearrangements, sex identification and applied genetics.

Cichlid fishes have been the subject of increasing scientific interest because of their rapid adaptive radiation that has led to an extensive ecological diversity and because of their enormous importance to tropical and subtropical aquaculture. Most cichlid species, especially the South American ones, are poorly known, and most of the genetic/genomic information concerning cichlids has been derived from the Nile tilapia *Oreochromis niloticus*. Despite the morphological and ecological diversity of cichlid fish, most cytogenetic data, which are based upon the karyotype information from 135 species, only refer to the determination of haploid/diploid number. The African cichlids have a modal number of 44 chromosomes, whereas the Neotropical cichlids have 48 diploid chromosomes [Feldberg et al., 2003]. The genus *Cichla* (peacock bass or tucunaré) is a widely distributed, endemic Neotropical cichlid and consists of 15 species [Kullander and Ferreira, 2006] of great importance as a food resource and to sport fishing. Moreover, phylogenetic data have shown that this genus is the sister clade of all other groups in the Cichlinae sub-family (Neotropical clade) [Smith et al., 2008]. Despite the importance of *Cichla* species, only basic cytogenetic data, including chromosome number and morphology and the distribution of heterochromatin and NOR (nucleolus organizer regions) sites, have been published for 2 species of the genus [Thompson, 1979; Brinn et al., 2004]. The *Cichla* species have 48 diploid chromosomes in accordance with the modal chromosome number for the South American cichlids, but the chromosomal formulas differ from the rest of the cichlids due to the presence of only subtelo-acrocentric chromosomes in the karyotype.

To further our understanding on the organization of cichlid genomes, we have isolated and characterized several classes of repeated DNA elements from the genome of *Cichla kelberi*. Since some African cichlid species genomes are being completely sequenced [The International Cichlid Genome Consortium, 2006], it will be of particular interest to investigate the genome structure of South American cichlid species for future comparative analyses. For that reason, the nucleotide sequence and chromosomal distribution of repeated DNA elements were investigated in *C. kelberi*, which led to the discovery of new repeated elements and also contributed to our knowledge about fish genome organization.

Material and Methods

Animals, DNA Samples, Chromosome Preparation and Banding

Animal samples were collected from the Araguaia River (São Felix do Araguaia, MT, Brazil), according to Brazilian laws for environmental protection (wild collection permit, SISBIO/15729-1). Tissue samples were collected from the animals and stored in 100% ethanol, and the genomic DNA was extracted using standard phenol-chloroform procedures [Sambrook and Russel, 2001]. Metaphase chromosome spreads were prepared from anterior kidney cells with in vivo colchicine treatment [Bertollo et al., 1978]. The constitutive heterochromatin was detected using saline solution [Sumner, 1972], and NORs were stained with silver nitrate [Howell and Black, 1980].

Isolation of Repeated DNA Elements

Repeated DNA elements were isolated via PCR (Polymerase Chain Reaction) and restriction enzyme digestion of genomic DNA. The first class of repeated DNA elements that were isolated by PCR was the ribosomal RNA (rRNA) genes. Copies of the 18S rRNA gene were amplified with the primers 18Sf (5′-CCG CTT TGG TGA CTC TTG AT) and 18Sr (5′-CCG AGG ACC TCA CTA AAC CA), which were designed based upon the sequence of the catfish *Ictalurus punctatus* (GenBank accession number AF021880) to amplify an approximately 1,400 bp DNA segment of the 18S rRNA gene. The 5S rRNA genes were isolated with the primers 5Sf (5′-TAC GCC CGA TCT CGT CCG ATC) and 5Sr (5′-CAG GCT GGT ATG GCC GTA AGC), which were designed from the 5S rRNA gene sequence from *Salmo gardnerii* [Komiya and Takemura, 1979] to amplify complete repeats of 5S rDNA [Martins and Galetti Jr., 1999].

The second repeated DNA class that was analyzed corresponds to the transposable element *Tc1* and to the retroelements *Rex1*, *Rex3* and *Rex6*. They were isolated by PCR with the primer sets, as follows: the *Tc1* element primer Tc1 (5′-TAC AGT GCC TTG CAT AAG TAT TCA CC), which anneals to the inverse repeats that flank the transposase gene of the element [Leaver, 2001]; the *Rex1* element primers Rex1f (5′-TTC TCC AGT GCC TTC AAC ACC) and Rex1r (5′-TCC CTC AGC AGA AAG AGT CTG CTC) [Volff et al., 2000]; the *Rex3* element primers Rex3f (5′-CGG TGA YAA AGG GCA GCC CTG) and Rex3r (5′-TGG CAG ACN GGG GTG GTG GT) (Volff et al., 1999); and the *Rex6* element primers Rex6f (5′-TAA AGC ATA CAT GGA GCG CCAC) and Rex6r (5′-GGT CCT CTA CCA GAG GCC TGGG) [Volff et al., 2001].

Repeated DNA elements were also isolated using restriction enzyme digestion of whole genomic DNA. DNA samples (8 μg) were digested to completion with the endonucleases *Afa*I, *Alu*I, *Bcl*I, *Bgl*II, *Dra*I, *Eco*RI, *Eco*RV, *Hae*III, *Hin*dIII, *Msp*I, *Ssp*I and *Xba*I and then electrophoresed in a 1% agarose gel. The endonuclease *Xba*I generated a band of approximately 650 bp, which was a candidate to contain repeated DNA elements, and that band was purified from the agarose gel for cloning.

Cloning, Sequencing and Sequence Analysis of Repeated DNAs

The PCR-generated amplicons of the transposon elements *Tc1*, *Rex1*, *Rex3* and *Rex6* were inserted in the plasmid pGEM-T (Promega) and the *Xba*I-generated DNA fragments were inserted in the plasmid vector pMOS (GE Healthcare Life Sciences). Ligation

products were transformed into DH5α *Escherichia coli* competent cells. Positive recombinant clones were recovered and stored for future analysis. The positive clones were sequenced on an ABI Prism 3100 DNA sequencer (Perkin-Elmer) with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer), and the sequences that were obtained were deposited in the NCBI database under the accession numbers FJ687584–FJ687590. The sequences were used for Blastn [Altschul et al., 1990] searches at the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/blast) to check for any similarity of the isolated sequences to any known sequences that have been deposited in the nucleotide collection (nt/nr), whole-genome shotgun reads (WGS), genomic survey sequences (GSS) and high throughput genomic sequences (HTGS) databases of GenBank.

The sequences recovered from GenBank and the sequences obtained here were aligned with the software BioEdit 7.0.9 [Hall, 1999], and converted to PAUP format using the software DAMBE [Xia and Xie, 2001]. Evolutionary models of each group of sequences (i.e. *Tc1*, *Rex3*, *Rex6* and *RCk*) were determined using a hierarchical hypothesis test of alternative models implemented with Modeltest 3.7 [Posada and Crandall, 1998] together with software PAUP* beta version 4.0b10 [Swofford, 2002]. The nucleotide substitution model based on Hasegawa, Kishino and Yano (HKY) [Hasegawa et al., 1985] was suitable for *RCk*, *Rex6* and *Tc1* sequences; Kimura 80 (K2P) [Kimura, 1980] fit *Rex1* sequences and General time reversible (GTR) [Tavaré, 1986] properly match *Rex3* data. Models incorporating rate variation (G) and PINVAR with 4 G-distributed rate classes [Swofford et al., 1996] were utilized for all likelihood analyses. The Bayesian-likelihood method of phylogenetic analysis [Huelsenbeck et al., 2001] was applied to evaluate alternative tree topologies using MrBayes v.3.0 [Ronquist and Huelsenbeck, 2003]. A total of 1,000,000 generations sampled at every 100th generation were tested by 4 chains running at once. Majority-rule consensus trees were produced with the software TreeExplorer implemented in MEGA 4 [Tamura et al., 2007].

Chromosome in situ Hybridization

Mitotic chromosomes were submitted to FISH (Fluorescence in situ hybridization) [Pinkel et al., 1986] using the PCR products from the 5S and 18S rRNA genes and clones containing the repetitive elements *Tc1*, *Rex1*, *Rex3*, *Rex6* and *RCk* as probes. The probes were labeled by nick translation with biotin-14-dATP for separate hybridization and with biotin-14-dATP (18S probe) and digoxigenin (5S probe) for simultaneous hybridization. The metaphase chromosome slides were incubated with RNase (40 μg/ml) for 1.5 h at 37°C. After the chromosomal DNA was denatured in 70% formamide, 2× SSC for 4 min at 70°C, the hybridization mixtures, which contained 100 ng of the denatured probe, 10 mg/ml dextran sulfate, 2× SSC and 50% formamide in a final volume of 30 μl, were dropped on the slides, and the hybridization was performed overnight at 37°C in a 2× SSC moist chamber. Post-hybridization washes were carried out at 37°C in 2× SSC, 50% formamide for 15 min, followed by a second wash in 2× SSC for 15 min and a final wash at room temperature in 4× SSC for 15 min. Detection of the biotin- and digoxigenin-labeled probes was carried out with Avidin-FITC (Sigma) and Anti-Dig-Rhodamine (Roche), respectively. Chromosomes were counterstained with propidium iodide (0.2%) or DAPI (0.01 mg/ml) diluted in antifade (Vector). Hybridized chromosomes were analyzed using an Olympus BX 61 microscope, and the images were captured with the Olympus DP70 digital camera with the software Image-Pro MC 6.0. Karyotypes were arranged in order of decreasing chromosome size. The extent of the hybridized signals was estimated as a percentage of the whole chromosomes using the software Image Tool.

Results and Discussion

PCR and Nucleotide Sequence Analysis of Repeated DNAs

The isolated 18S rRNA gene fragment contains an ~1,400-bp internal segment of the gene. The 18S primers were successfully applied in several non-related fish species (data not shown), which demonstrated their applicability for the amplification of this gene for use as a probe for FISH. The physical chromosome mapping of the DNA fragment that was isolated with the primers 18Sf and 18Sr coincided with the NOR region, which confirmed that the PCR-isolated DNA segment corresponded to the 18S rRNA gene. The 5S rDNA repeats of ~230 bp were recovered by PCR with the primers 5Sf and 5Sr and were also used for FISH.

The PCR of the transposon *Tc1* element generated an electrophoretic DNA band of approximately 370 bp, which was cloned and sequenced. The nucleotide sequence of 4 clones was determined, and the consensus sequence revealed a 369-bp DNA fragment that was flanked by inverted repeats with similarity to the *Tc1* element and other dispersed sequences in several vertebrate genomes, including fish and amphibians. The *Tc1*-like isolated sequence corresponds to the 5′ flanking sequence of the transposase-coding region of the element [Leaver, 2001]. The *Tc1* transposon belongs to the *Tc1-mariner* superfamily and is exceptionally widespread among living organisms from protozoa to vertebrates [Ivics et al., 1996]. *Tc1* elements were first identified in the invertebrates *Caenorhabditis elegans* and *Drosophila* sp., and they share several structural similarities, such as the presence of long terminal inverted repeats (TIRs) with an identical 5/6-bp sequence at or near the end and an open reading frame that encodes a 347-amino acid transposase [Brezinsky et al., 1990; Anderson et al., 1992; Avancini et al., 1996].

An NCBI database search revealed the identification of *Tc1*-like sequences among several fish and amphibian species, and these were comparatively analyzed (fig. 1). The *Tc1*-like sequence of *C. kelberi* clustered with sequences of *Takifugu* (Tetraodontiformes), *Salmo* (Salmoniformes) and *Gasterosteus* (Gasterosteiformes) in a

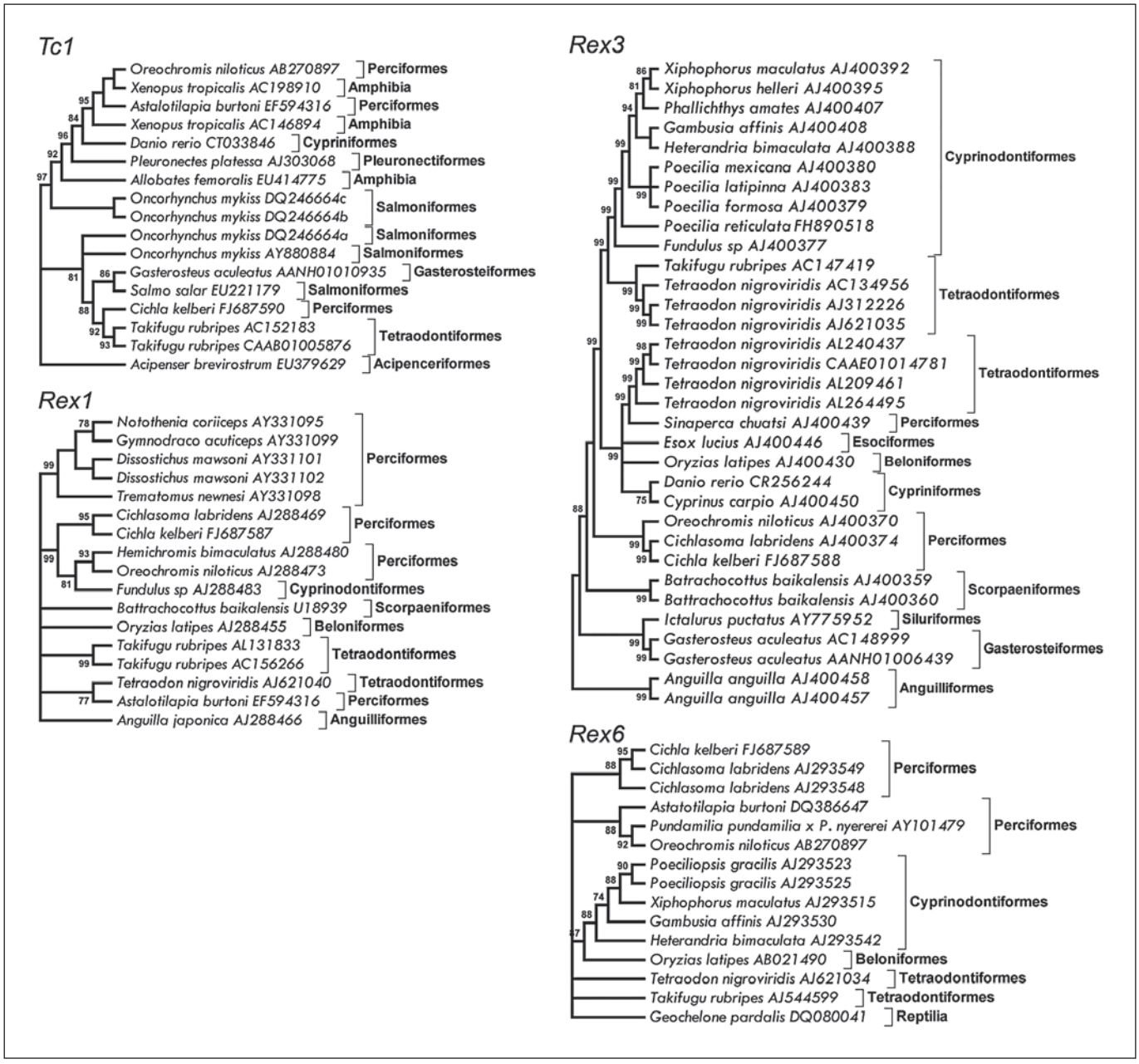

Fig. 1. Maximum likelihood trees of *Tc1*, *Rex1*, *Rex3* and *Rex6* repeated elements among fishes and other vertebrates as obtained by Bayesian inference. The species names and NCBI entries are indicated. Bootstrap values above 70 are shown.

separate branch that contained the cichlids *Oreochromis niloticus* and *Astatotilapia burtoni*. The *Tc1* element has a distribution among the species that is inconsistent with the phylogenetic relationship of different fish groups and also with relationship of these fishes to amphibians. Such a result suggests that events of horizontal transmission could have occurred during the evolutionary history of the *Tc1* element in vertebrates, as has been previously suggested for some organisms [Ivics et al., 1996; Leaver, 2001]. It seems most probable that the *Tc1*-like sequence isolated here might be a product of miniature inverted-repeat transposable elements (MITEs). These elements are short transposons (100–600 bp), which are distinguished from other non-autonomous elements by high

copy numbers and length homogeneity, that are thought to have arisen by amplification of a single or a few progenitor copies. Presumably, the progenitor copy arose via deletion of a larger transposon during gap repair. In many cases, sequence similarity between MITEs and the closest autonomous element is restricted to the TIRs [Feschotte et al., 2002; Feschotte and Pritham, 2007]. Some MITEs have shown TIRs and target-site duplications (TSD) that are similar to those of the *Tc1-mariner* superfamily, which supports the idea that these particular MITEs might have originated from this family of transposons [Freschotte et al., 2002]. Recently, a strong similarity between the TIRs, the target-site preference of a MITE and a *Tc1*-like transposon family was identified in the genome of a fungus species [Feschotte and Pritham, 2007].

The short length, lack of a transposase gene and the high similarity with the TIRs of other *Tc1* transposons of fish species suggest that the sequence isolated in this work is more closely related to a MITE than a *Tc1* transposon. The *Cichla* genome probably contains both *Tc1*-like transposons and MITEs that are related to this transposon, but the primer might have amplified more MITEs than the full length transposable element.

The retrotransposons *Rex1, Rex3* and *Rex6* generated an electrophoretic DNA band of approximately 600 bp. The primer sets that were employed were very efficient in the amplification of these DNA fragments by PCR and have been employed in the isolation of these repeated elements in different fish groups [Capriglione et al., 2002; Ozouf-Costaz et al., 2004]. The PCR fragments obtained here were cloned and the nucleotide sequences were determined for 3 clones of each element that were aligned and a consensus sequence generated.

In general, *Rex1* sequences of *C. kelberi* showed similarity higher than 80% to *Rex1* elements that were previously identified in other vertebrates and also to other dispersed sequences in the genome of the Perciformes fish. The PCR-isolated *Rex1* segment corresponds to the encoding domains 3–7 of the reverse transcriptase (RT) gene [Volff et al., 2000]. Relevant similarity of *Rex1* to the fish orders Anguilliformes, Perciformes, Beloniformes, Cyprinodontiformes and Tetraodontiformes was detected (fig. 1). A detailed analysis of the *Rex1* element was previously conducted among fish species, which demonstrated the presence of multiple ancient lineages that underwent several independent and recent bursts of retrotransposition within these fish genomes [Volff et al., 2000].

The obtained nucleotide segment of *Rex3* transposon corresponds to the encoding domain 1, 2, 2A, A and B of the RT gene [Volff et al., 1999]. The *C. kelberi Rex3* element contains sequences that have 85–92% similarity to the retrotransposon *Rex3* of the cichlids *Cichlasoma labridens* and *O. niloticus* and lower similarities (77–82%) to elements that are found in the teleost species of Beloniformes, Cyprinodontiformes, Cypriniformes, Esociformes, non-cichlid Perciformes, Scorpaeniformes and Tetraodontiformes (fig. 1). Previous studies have shown that this element is widely distributed within fish genomes [Volff et al., 2001; Ozouf-Costaz et al., 2004]. Cichlid species are organized in the same branch of the tree and the South American cichlids *Cichlasoma labridens* and *Cichla kelberi* are organized in a branch that is distinct from that of the African cichlid *O. niloticus* (fig. 1). Segments of the *Rex3* elements were previously isolated from the South American cichlid *Astronotus ocellatus* [Mazzuchelli and Martins, 2009], but these fragments corresponded to the 5′ flanking domain of the *Rex3* element and do not correspond to the segment that was characterized for *C. kelberi*.

The obtained *Rex6* sequences of *C. kelberi* correspond to the C-terminal part of the restriction enzyme-like endonuclease of the retrotransposon element [Volff et al., 2001]. The *Rex6* element has been poorly investigated in vertebrates, and there are few nucleotide sequences available through NCBI. Therefore, our analysis will contribute additional information about the *Rex6* element in the fish genome. The *Rex6* element that was identified from the *Cichla* genome showed high similarity with related *Rex6* elements from other Cichlidae species *(C. labridens*, *O. niloticus* and *A. burtoni)* and from the Tetraodontiformes, Cyprinodontiformes and Beloniformes orders (fig. 1). The retrotransposon *Rex6* was active during teleost evolution and successfully invaded the genomes of several fish species. This retrotransposon was also identified in the genome of the reptilian *Geochelone pardalis* (fig. 1).

The restriction DNA fragments that were generated by the enzyme *Xba*I were cloned, and the nucleotide sequence analysis of 5 positive clones revealed that 3 of the isolated sequences possessed a high nucleotide similarity to dispersed sequences in the genome of several fishes and 2 amphibian species. The most noticeable characteristic of this element was its similarity to flanking sequences of the RT gene of the fish *Oryzias melastigma* (NCBI accession number DQ286655). The similarity to dispersed sequences and to a transcriptase reverse gene suggested that the isolated sequence corresponded to a retrotransposon that we called *RCk* (Retrotransposon of *C. kelberi*) element. Although comparative analyses of the complete

RCk element of *C. kelberi* and *O. melastigma* TR sequence showed a low similarity (35–38%), a higher similarity (76–82%) was detected in a segment that corresponded to an internal sequence of 273 nucleotides. A Blast search with this *RCk* element identified 70–87% similarity to sequences from fish species that represent the orders Perciformes, Salmoniformes, Cypriniformes, Cyprinodontiformes, Gasterosteiformes, Beloniformes and Polypteriformes and 74–82% similarity to sequences from the amphibians *Xenopus tropicalis* and *Rana chensinensis* (fig. 2). The presence of an *RCk* element in fish and amphibians suggests that the origin of this element occurred early in the evolutionary history of vertebrates, before the divergence of the major groups. The *RCk* element seems to have several ancient lineages that gave origin to the diversity of sequences detected among fish orders. The *RCk* element was detected in the cichlid species *M. conophoros*, *L. fuelleborne*, *M. zebra* and *A. burtoni* (fig. 2).

Although fish genomes are more compact compared to those of mammals, a higher diversity among their retrotransposable elements is characteristic of their genomes when compared to the human and mouse genomes [Volff et al., 2003]. Given this, investigation of fish repeated DNA elements can contribute to the knowledge of their genomes, especially since these repeated elements remain as gaps even within those genomes that have been reported to be completely sequenced. These gaps exist due to the difficulty in the correct identification of the position, array and repeat number of these repeated DNA elements. In addition, some specific repeated DNA elements could also be applied as molecular markers to track the evolutionary history of particular clades.

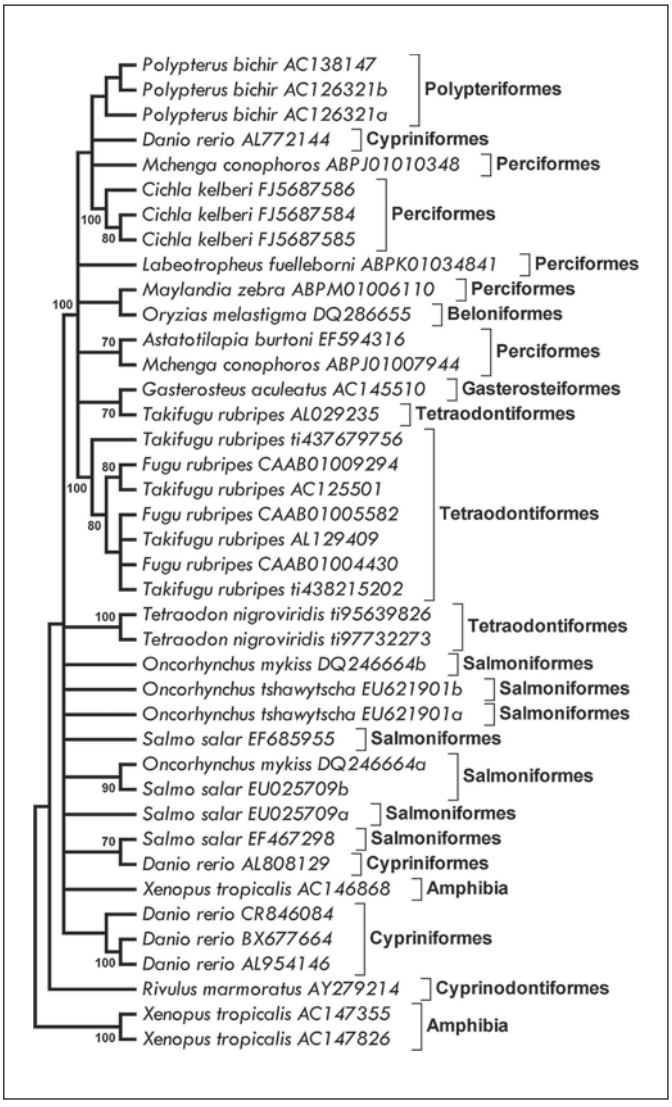

Fig. 2. Maximum likelihood tree of *RCk* element among fishes and other vertebrates as obtained by Bayesian inference. The species names and NCBI entries are indicated. Bootstrap values above 70 are shown.

Basic Cytogenetic Analysis

Metaphase chromosomes were obtained from 8 males and 5 females of *C. kelberi*. The presence of 48 subtelo-acrocentric (st/a) chromosomes was constant between the sexes (fig. 3a). The NOR regions were conserved in the terminal position of the long arm of chromosome pair 2, which contains a large secondary constriction (fig. 3a) and a strong mark that is visible by silver nitrate staining (fig. 3c). The heterochromatin was distributed mainly in the centromeric areas of most chromosomes in the complement (fig. 3b). Heterochromatic segments were also observed in coincidence with the NOR region and in an interstitial position on the long arm of chromosome pair 2, which is the NOR-bearing chromosome. The chromosome morphology and the heterochromatin distribution of *C. kelberi* are similar to that of *C. monoculus* and *C. temensis* from the Amazon region [Brinn et al., 2004]. The NOR location in the second pair was also observed in the *C. monoculus* karyotype, while this chromosomal marker occurs in the third pair in *C. temensis*, which indicates a variability of this character in the genera. Although the chromosome number (2n = 48) has been maintained during the diversification of South American cichlids, the presence of only subtelo-acrocentric chromosomal elements in the karyotype is exclusive to *Cichla*. The remaining cichlids, including the African ones, possess karyotypes with the presence of meta- and submetacentric chromosomes. The karyotype formula 2n = 48

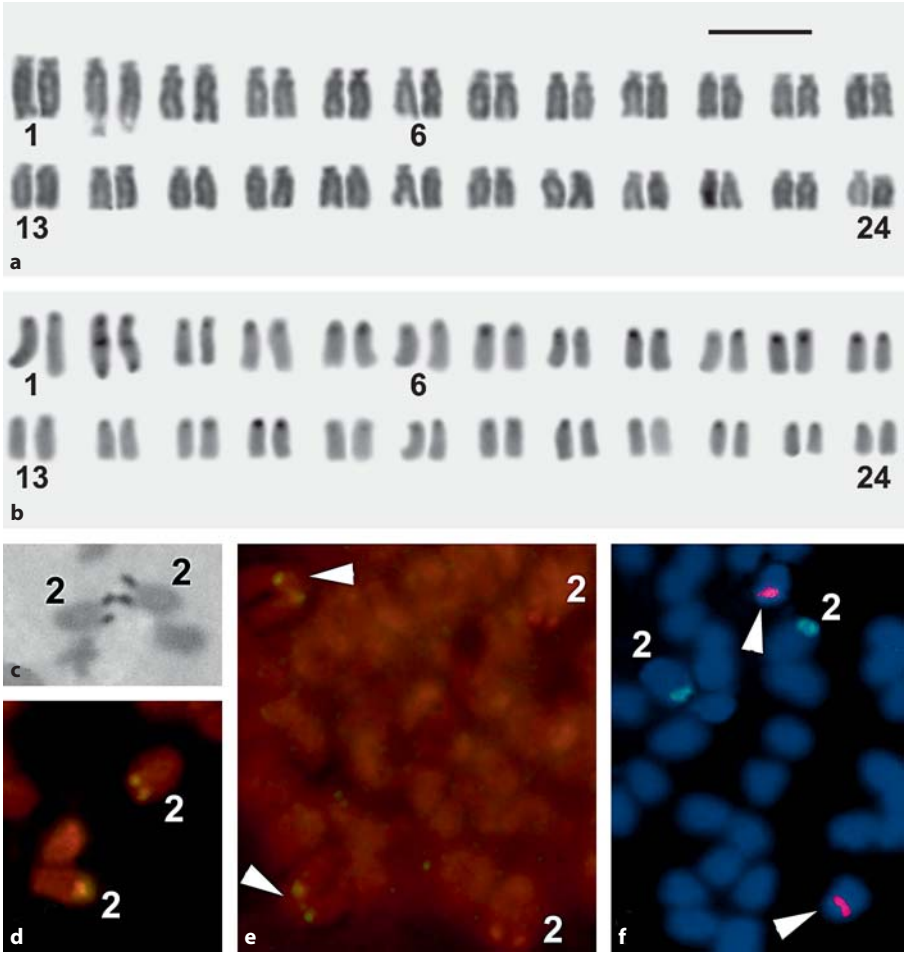

Fig. 3. Chromosomal analysis of *Cichla kelberi* by means of Giemsa staining (**a**), C-banding (**b**), Ag-NOR staining (**c**) and FISH with 18S rDNA (**d**) and 5S rDNA (**e**) probes. **f** Two-color FISH with 18S (green) and 5S (red) rDNA probes. The NOR chromosome pair (number 2) is indicated, and the arrows indicate the 5S sites (**e**, **f**). Scale bar indicates 5 μm.

st/a elements is characteristic of Perciformes, which includes the groups that are related to Cichlidae, such as Sciaenidae [Feldberg et al., 1999; Accioly and Molina, 2008], Pomacentridae [Molina and Galetti Jr., 2002] and Haemulidae [Nirchio et al., 2007]. Such data suggest that *Cichla* retains the ancestral karyotype of the group. The ancestral karyotype could have undergone major changes in its macro-structure, which could have led to the extensive karyotype diversification that is currently observed among cichlids. Such an observation is consistent with several proposed phylogenies for the family [Smith et al., 2008], which generally include the genus *Cichla* in the most basal clade of the group.

Chromosome Physical Mapping of Repeated DNA Sequences

Ribosomal RNA genes are among the most mapped sequences in the chromosomes of fishes. The 18S rRNA genes of *Cichla kelberi* mapped in coincidence to the NOR regions in the long arm of chromosome pair 2, whereas the 5S rRNA genes were located in an interstitial position of a different chromosome pair (fig. 3d, e). The different chromosomal location of the 18S and 5S rRNA genes was based upon the heterochromatin presence in the NOR region that was positively stained after propidium iodide counterstaining of 5S rDNA hybridized metaphases (fig. 3e) or after double FISH with 18S and 5S probes (fig. 3f). No intra- or inter-individual variation in the number of rRNA gene sites was observed, which reflects the conserved pattern of the chromosomal macro-structure of the genus *Cichla*.

Ribosomal RNA genes were previously mapped on the chromosomes of the cichlids *O. niloticus*, *Geophagus brasiliensis* and *Cichlasoma facetum*. The 18S rRNA genes were located in the terminal position of the short arms of 6 chromosomes in *O. niloticus* [Martins et al., 2000] and in the short arms of one chromosome pair in *G. brasiliensis* and *C. facetum* [Vicari et al., 2006]. The 5S rRNA genes

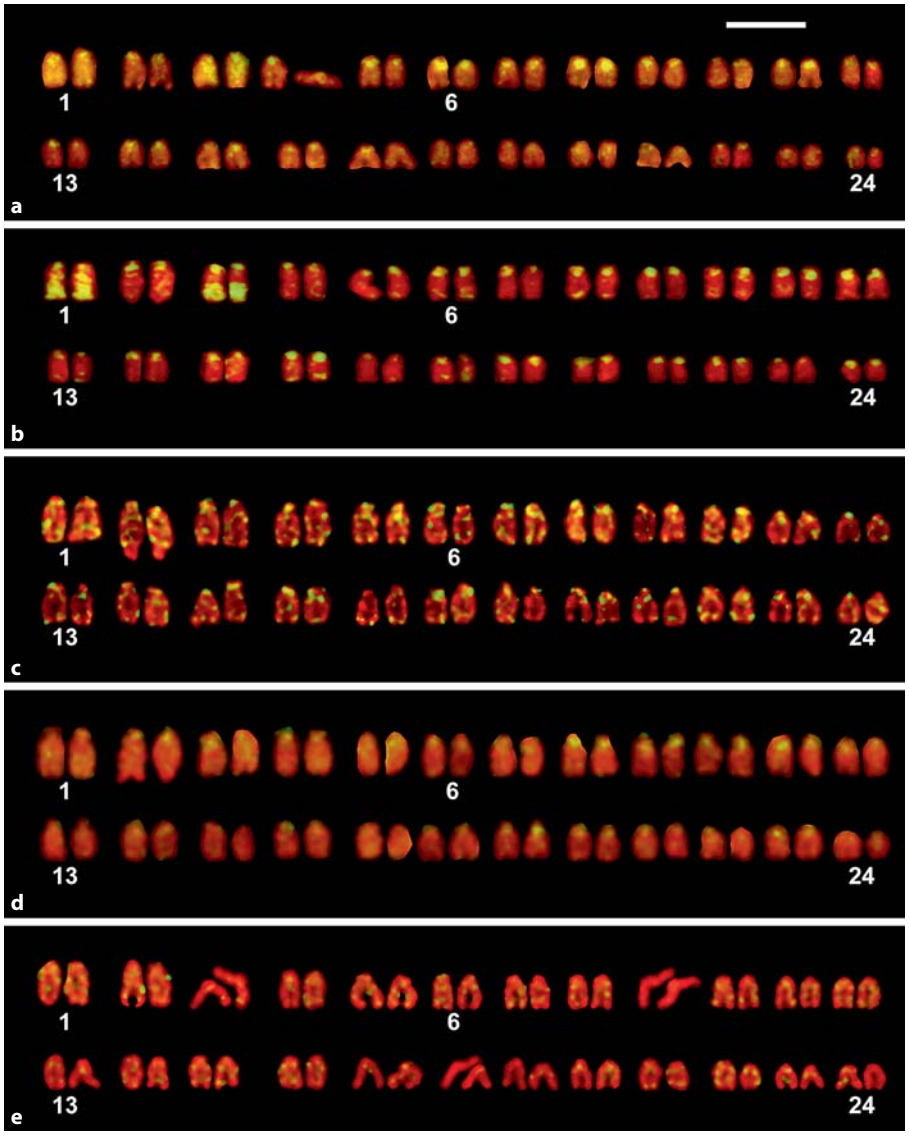

Fig. 4. Distribution of the repeated elements *Rex1* (**a**), *Rex3* (**b**), *Rex6* (**c**), *Tc1* (**d**) and *RCk* (**e**) on *Cichla kelberi* chromosomes. Scale bar indicates 5 μm.

were detected in 6 chromosomes of *O. niloticus*, being interstitially located on 2 chromosome pairs and in the terminal position of the short arm of one chromosome pair [Martins et al., 2002]. Only one chromosome pair of the species *G. brasiliensis* and *C. facetum* had an interstitial 5S site [Vicari et al., 2006]. Although the rRNA genes have been mapped in only a few cichlids, it is interesting to note that the African species *O. niloticus* possesses multiple 18S and 5S sites, whereas the South American species *(G. brasiliensis, C. facetum* and *C. kelberi)* possess only one chromosomal site for each gene class. Another interesting characteristic related to the chromosomal position of the rRNA genes sites is the fact that the 5S sites are located in an interstitial position within the cichlid chromosomes, with the exception of the one chromosomal site for *O. niloticus,* and the 18S sites were always located in terminal positions. These characteristics were also observed for several other fishes, which suggest that such a chromosomal pattern for the distribution of rRNA genes seems to be a trend in this vertebrate group.

The repeated elements *Rex1*, *Rex3*, *Rex6* and *Tc1* were clustered predominantly in the centromere in coincidence with heterochromatic areas and also as small dispersed signals along most chromosomes (fig. 4). Although transposable elements are abundant at most centromeres and their flanking regions, the significance of such a distribution is not understood [Wong and Choo, 2004]. In addition to the compartmentalization within

the centromeric heterochromatin, each element has a particular distribution pattern. The preferential occurrence of the retroelements in the centromeric heterochromatic region of *C. kelberi* chromosomes is a common feature that is observed among eukaryotes. This characteristic could be related to the involvement of repeated DNA elements with centromeric function [Dawe, 2003] or to the lower selective pressure that acts against repeated elements [Eickbush and Furano, 2002]. On the other hand, the repeated elements are also expected to accumulate in heterochromatin that is characterized by lower gene density and reduced recombination [Szauter, 1984]. Heterochromatin can reduce the probability of the elimination of inserted sequences by ectopic recombination, and this mechanism is believed to play a major role in controlling the number of transposable element copies [Charlesworth and Langley, 1989]. The negative selection that acts against the accumulation of repeated sequences in generich euchromatin may also have an effect on the distribution of repeated elements [Eickbush and Furano, 2002; Volff et al., 2003]. Similar patterns of repeated DNA organization have already been described in the African cichlid *Oreochromis niloticus* [Ferreira and Martins, 2008], in the South American cichlid *Astronotus ocellatus* [Mazzuchelli and Martins, 2009], in other non-cichlid fishes [Ozouf-Costaz et al., 2004; Herrán et al., 2008] and in other animals [Mravinac et al., 2004; Alkan et al., 2007].

Rex1 and *Rex3* have similar chromosomal distributions as they are both centromeric and are both spread on the long euchromatic arms of the first and third chromosome pairs (fig. 4a, b). These elements occupy a considerable extent of the euchromatin of chromosomes 1 and 3 (table 1). Based upon observations in *Drosophila*, it has been hypothesized that the differences in the euchromatic/heterochromatic distribution of retrotransposons could be the outcome of an initial loss of euchromatic elements that is followed by a reinvasion of transposons [Busseau et al., 1994]. The similar distribution of the *Rex1* and *Rex3* elements suggests that these sequences could be evolving together, have cooperative activity within the *C. kelberi* genome and be related with regard to structural or functional aspects. The compartmentalization of *Rex1* and *Rex3* elements in the centromeres and heterochromatin was also observed in the fish *Notothenia coriiceps* [Ozouf-Costaz et al., 2004]. The results observed for the distribution of *Rex1* and *Rex3* in the chromosomes of *C. kelberi,* however, direct a new approach to the study of the genomic organization of transposons. Considering the existence of a correlation between karyotype rearrangement and retrotransposon activity [Ozouf-Costaz et al., 2004], we could speculate that chromosomes 1 and 3 of *C. kelberi* have undergone rearrangement events during their evolutionary history. On the other hand, the repeated elements *Rex1* and *Rex3* could be accumulating in chromosomes 1 and 3 as a consequence of recombination suppression.

Table 1. Distribution, in percentage of chromatin extent, of the transposable elements *Rex1* and *Rex3* on the chromosomes of *C. kelberi*.

	Rex1, %	*Rex3*, %
Entire chromosome set	39.6	33.9
Pericentromeric regions	33.3	28.1
Euchromatic areas	6.3	5.8
First chromosome pair	62.8	67.8
Third chromosome pair	66.4	47.03

To the best of our knowledge, this report is the first description of the chromosomal physical mapping of retroelement *Rex6* in a vertebrate species. *Rex6* shows a diffuse pattern of localization and is not only compartmentalized to centromeric heterochromatin, but is also spread along the chromosomes of *C. kelberi* (fig. 4c). This indicates that *Rex6* localization is governed by distinct evolutionary mechanisms and can evolve independently of other repeated sequences.

Chromosomal mapping data on *Tc1*-like sequences in vertebrates is scarce, with these elements mapped in the chromosomes of the Antarctic ice-fish *Chionodraco hamatus* [Capriglione et al., 2002], the pufferfish *Tetraodon nigroviridis* [DaSilva et al., 2002] and the frog *Rana sculenta* [Pontecorvo et al., 2000]. In *C. kelberi*, this transposable element is observed in centromeres and dispersed along the chromosomal arms of most chromosomes of the complement (fig. 4d). This pattern is similar to the results that were observed in the other analyzed vertebrates. In *R. sculenta*, a sequence that originated from a *Tc1*-like element was organized in tandem arrays in the centromeres of few chromosomes [Pontecorvo et al., 2000]. Another example of such remarkable compartmentalization was observed in *T. nigroviridis*, where the *Tc1*-like elements were clustered in the heterochromatic short arms of 6 subtelocentric chromosome pairs [DaSilva et al., 2002].

The *RCk* transposable element had almost the same hybridization pattern as *Rex6*, with the exception of fewer signals in the pericentromeric region (fig. 4e). The

weak hybridization of *RCk* might be related to the presence of small copy numbers of this retrotransposon within the genome of *Cichla*. Given that *RCk* is distributed overall among fishes, however, such an element should be better investigated in other cichlids, at the very least. Compared to the other repeated elements that were investigated here, *RCk* was not compartmentalized in heterochromatin or in the centromeric areas, which suggests that different repeated element classes are governed by distinct mechanisms of evolution.

Conclusions

Considering that the transposable elements *Tc1*, *Rex1*, *Rex3*, *Rex6* and *RCk* are widespread among fishes, their application as physical chromosome markers appears to hold great promise to contribute to our knowledge of fish cytogenetics. The integration of DNA sequences with physical chromosome mapping of repeated DNA elements can provide a better landscape of the genome, which is not yet clearly defined even in those completely sequenced genomes. Therefore, a complete understanding of the relationship between chromosome structure and function requires an understanding of the organization of the repeated portions of the genome. Given that some African cichlid genomes are being completely sequenced [The International Cichlid Genome Consortium, 2006], the genome structure of South American cichlid species is of particular interest for comparative analysis. Although cytogenetic studies have been carried out on a large number of fish species in the last 2 decades, such analyses were mainly focused on basic karyotype structure, and very little work has been performed on the organization of DNA sequences within chromosomes.

Acknowledgements

The authors are grateful to FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), FAPEMAT (Fundação de Amparo à Pesquisa do Estado de Mato Grosso), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for financial support, and to Entre Rios Farm for assistance during fish sampling in the Araguaia river.

References

Accioly IV, Molina WF: Cytogenetic studies in Brazilian marine Sciaenidae and Sparidae fishes (Perciformes). Genet Mol Res 7:358–370 (2008).

Alkan C, Ventura M, Archidiacono N, Rocchi M, Sahinalp SC, et al: Organization and evolution of primate centromeric DNA from whole-genome shotgun sequence data. PLoS Comput Biol 3:181 (2007).

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. J Mol Biol 215:403–410 (1990).

Anderson P, Emmons SW, Moerman DG: Discovery of Tc1 in the nematode *Caenorhabditis elegans*, in Fedoroff N, Botstein D (eds): The Dynamic Genome, pp 319–333 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1992).

Avancini RMP, Walden KKO, Robertson HM: The genomes of most animals have multiple members of the Tc1 family of transposable elements. Genetica 98:131–140 (1996).

Bertollo LAC, Takahashi CS, Moreira-Filho O: Cytotaxonomic consideration on *Hoplias lacerdae* (Pisces, Erythrinidae). Braz J Genet 1:103–120 (1978).

Brenner S, Elgar G, Sandford R, Macrae A, Venkatesh B, Aparício S: Characterization of the pufferfish *(Fugu)* genome as a compact model vertebrate genome. Nature 366:265–268 (1993).

Brezinsky L, Wang GVL, Humphreys T, Hunt J: The transposable element Uhu from Hawaiian *Drosophila* – member of the widely dispersed class of Tc1-like transposons. Nucleic Acids Res 18:2053–2059 (1990).

Brinn MNA, Porto JIR, Feldeberg E: Karyological evidence for interspecific hybridization between *Cichla monoculus* and *C. temensis* (Perciformes, Cichlidae) in the Amazon. Hereditas 141:252–257 (2004).

Busseau I, Chaboissier MC, Pélisson A, Bucheton A: I factors in *Drosophila melanogaster*: transposition under control. Genetica 93: 101–116 (1994).

Capriglione T, Odierna G, Caputo V, Canapa A, Olmo E: Characterization of a Tc1-like transposon in the Antarctic ice-fish, *Chionodraco hamatus*. Gene 295:193–198 (2002).

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

Charlesworth B, Langley CH: The population genetics of *Drosophila* transposable elements. Annu Rev Genet 23:251–287 (1989).

DaSilva C, Hadji H, Ozouf-Costaz C, Nicaud S, Jaillon O, et al: Remarkable compartmentalization of transposable elements and pseudogenes in the heterochromatin of the *Tetraodon nigroviridis* genome. Proc Natl Acad Sci USA 99:13636–13641 (2002).

Dawe RK: RNA interference, transposons, and the centromere. Plant Cell 15:297–301 (2003).

Eickbush TH, Furano AV: Fruit flies and humans respond differently to retrotransposons. Curr Opin Genet Dev 12:669–674 (2002).

Feldberg E, Porto JIR, Santos EBP, Valentim FCS: Cytogenetics studies of two freshwater sciaenids of the genus *Plagioscion* (Perciformes, Scianidae) from the Central Amazon. Genet Mol Biol 22:351–356 (1999).

Feldberg E, Porto JIR, Bertollo LAC: Chromosomal changes and adaptation of cichlid fishes during evolution, in Val AL, Kappor BG (eds): Fish Adaptations, pp 285–308 (Science Publishers, New Dehli 2003).

Ferreira IA, Martins C: Physical chromosome mapping of repetitive DNA sequences in Nile tilapia *Oreochromis niloticus*: Evidences for a differential distribution of repetitive elements in the sex chromosomes. Micron 39: 411–418 (2008).

Feschotte C, Pritham EJ: DNA transposons and the evolution of eukaryotic genomes. Ann Rev Genet 41:331–368 (2007).

Feschotte C, Jiang N, Wessler SR: Plant transposable elements: Where genetics meets genomics. Nat Rev Genet 3:329–341 (2002).

Grady DL, Ratliff RL, Robinson DL, McCanlies EC, Meyne J, Moyzis RK: Highly conserved repetitive DNA sequences are present at human centromeres. Proc Natl Acad Sci USA 89:1695–1699 (1992).

Hall TA: BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95–98 (1999).
Hasegawa M, Kishino K, Yano T: Dating the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174 (1985).
Herrán De la R, Robles F, Navas JI, López-Flores I, Herrera M, et al: The centromeric satellite of the wedge sole (*Dicologoglossa cuneata*, Pleuronectiformes) is composed mainly of a sequence motif conserved in other vertebrate centromeric DNAs. Cytogenet Genome Res 121:271–276 (2008).
Horvath JE, Bailey JA, Locke DP, Eichler EE: Lessons from the human genome: transitions between euchromatin and heterochromatin. Hum Mol Genet 10:2215–2223 (2001).
Howell WM, Black DA: Controlled silver staining of nucleolus organizer regions with a protective colloidal developer: a 1-step method. Experientia 36:1014–1015 (1980).
Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP: Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294: 2310–2314 (2001).
Ivics Z, Izsvak Z, Minter A, Hackett PB: Identification of functional domains and evolution of Tc1-like transposable elements. Proc Natl Acad Sci USA 93:5008–5013 (1996).
Kimura M: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120 (1980).
Komiya H, Takemura S: Nucleotide sequence of 5S ribosomal RNA from rainbow trout *(Salmo gairdnerii)* liver. J Biochem 86:1067–1080 (1979).
Kullander SO, Ferreira EJG: A review of the South American cichlid genus *Cichla*, with descriptions of nine new species (Teleostei: Cichlidae). Ichth Expl Fresh 17:289–398 (2006).
Leaver MJ: A family of *Tc1*-like transposons from the genomes of fishes and frogs: evidence for horizontal transmission. Gene 271: 203–214 (2001).
Martins C, Galetti PM Jr: Chromosome localization of 5S rRNA genes in *Leporinus* (Anostomidae, Characiformes). Chromosome Res 7:363–367 (1999).
Martins C, Wasko AP, Oliveira C, Wright J: Nucleotide sequence of 5S rDNA and localization of the ribosomal RNA genes to metaphase chromosomes of the Tilapiini cichlid fish, *Oreochromis niloticus*. Hereditas 133: 39–46 (2000).
Martins C, Wasko AP, Oliveira C, Porto-Foresti F, Parise-Maltempi PP, et al: Dynamics of 5S rDNA in the tilapia *(Oreochromis niloticus)* genome: repeat units, inverted sequences, pseudogenes and chromosome loci. Cytogenet Genome Res 98:78–85 (2002).
Mazzuchelli J, Martins C: Genomic organization of repetitive DNAs in the cichlid fish *Astronotus ocellatus*. Genetica 136:461–469 (2009).
Molina WF, Galetti PM Jr: Robertsonian rearrangements in the reef fish *Chromis* (Perciformes, Pomacentridae) involving chromosomes bearing 5S rRNA genes. Genet Mol Biol 25:373–377 (2002).
Mravinac B, Plohl M, Ugarkovic D: Conserved patterns in the evolution of *Tribolium* satellite DNAs. Gene 332:169–177 (2004).
Nirchio M, Gaviria JI, Oliveira C, Ferreira IA, Martins C: Cytogenetic analysis of three species of the genus *Haemulon* (Teleostei: Haemulinae) from Margarita Island, Venezuela. Genetica 131:135–140 (2007).
Ozouf-Costaz C, Brandt J, Körting C, Pisano E, Bonillo C, et al: Genome dynamics and chromosomal localization of the non-LTR retrotransposons Rex1 and Rex3 in Antarctic fish. Antarct Sci 16:51–57 (2004).
Pinkel D, Straume T, Gray JW: Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc Natl Acad Sci USA 83:2934–2938 (1986).
Pontecorvo G, Felice BD, Carfagna M: A novel repeated sequence DNA originated from a Tc1-like transposon in water green frog *Rana esculenta*. Gene 261:205–210 (2000).
Posada D, Crandall KA: Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818 (1998).
Ronquist F, Huelsenbeck JP: MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574 (2003).
Sambrook J, Russel DW: Molecular Cloning. A Laboratory Manual, 3rd ed (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2001).
Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard H: Genomic and genetic definition of a functional human centromere. Science 294:109–115 (2001).
Smith WL, Chakrabarty P, Sparks JS: Phylogeny, taxonomy, and evolution of Neotropical cichlids (Teleostei: Cichlidae: Cichlinae). Cladistics 24:625–641 (2008).
Sumner AT: A simple technique for demonstrating centromeric heterochromatin. Exp Cell Res 75:304–306 (1972).
Swofford DL: PAUP* – Phylogenetic analysis using parsimony (* and other methods). Version 4b10 (Sinauer, Sunderland 2002).
Swofford DL, Olsen GJ, Waddell PJ, Hillis DM: Phylogenetic inference, in Hillis DM, Mable BK, Moritz C (eds): Molecular Systematics, pp 407–514 (Sinauer, Sunderland 1996).
Szauter P: An analysis of regional constraints on exchange in *Drosophila melanogaster* using recombination-defective meiotic mutants. Genetics 106:45–71 (1984).
Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599 (2007).
Tavaré S: Some probabilistic and statisical problems on the analysis of DNA sequences. Lect Math Life Sci 17:57–86 (1986).
The International Cichlid Genome Consortium: Genetic basis of vertebrate diversity: the cichlid fish model. Available at http://hcgs.unh.edu/cichlid/ (2006).
Thompson KW: Cytotaxonomy of 41 species of Neotropical Cichlidae. Copeia 1979:679–691 (1979).
Vicari MR, Artoni RF, Moreira-Filho O, Bertollo LAC: Basic and molecular cytogenetics in freshwater Cichlidae (Osteichthyes, Perciformes). Karyotypic conservationism and divergence. Caryologia 59:260–266 (2006).
Volff JN, Körting C, Sweeney K, Schartl M: The non-LTR retrotransposon *Rex3* from the fish *Xiphophorus* is widespread among teleosts. Mol Biol Evol 16:1427–1438 (1999).
Volff JN, Körting C, Schartl M: Multiple lineages of the non-LTR retrotransposon *Rex1* with varying success in invading fish genomes. Mol Biol Evol 17:1673–1684 (2000).
Volff JN, Körting C, Froschauer A, Sweeney K, Schartl M: Non-LTR retrotransposons encoding a restriction enzyme-like endonuclease in vertebrates. J Mol Evol 52:351–360 (2001).
Volff JN, Bouneau L, Ozouf-Costaz C, Fischer C: Diversity of retrotransposable elements in compact pufferfish genomes. Trends Genet 19:674–678 (2003).
Wong LH, Choo KH: Evolutionary dynamics of transposable elements at the centromere. Trends Genet 20:611–616 (2004).
Xia X, Xie Z: DAMBE: Data analysis in molecular biology and evolution. J Hered 92:371–373 (2001).