Highly repetitive elements from Chinese bitterlings (genus *Rhodeus*, Cyprinidae)

Kenji Saitoh^{1*}, Takayoshi Ueda², Ryoichi Arai³,
Han-Ling Wu⁴, Sang-Rin Jeon⁵

¹Tohoku National Fisheries Research Institute, Shinhamma, Shiogama 985-0001,

²Department of Biology, Faculty of Education, Utsunomiya University,
Minemachi, Utsunomiya 321-0943,

³Department of Zoology, University Museum, University of Tokyo,
Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

⁴Ichthyological Division, Shanghai Fisheries University, 334
Jun Gong Road, Shanghai 200090, China and

⁵Sang Myung University, Hongji-dong, Jongno-gu,
Seoul 110, Korea

(Received 4 January 2000, accepted 20 October 2000)

We have isolated and characterized several highly repetitive DNA elements from two species of Chinese bitterlings, *Rhodeus atremius suigensis* and *R. ocellatus ocellatus*. They comprise a partly interspersed and partly tandem repetitive family of about 1.0 to 1.3kb in length. Individual elements showed considerable length variation, but genomic Southern blotting revealed two major length groups. Their restricted presence of these elements among related species and relative copy number differences indicated rapid change of genome structure in this group of fish. The isolated elements may be useful landmarks for further chromosomal studies.

INTRODUCTION

Eukaryotic genomes contain much redundancy. A considerable portion of the redundancy is due to repetitive DNA elements (Britten and Kohne 1968). Many of them do not have any effect on their host organisms' phenotype but propagate themselves within the host genomes (Doolittle and Sapienza 1980; Orgel and Crick 1980), while some have mutational activities due to insertion into functional genes (Hori et al. 1998; Izsvák et al. 1996). Their increase or decrease, therefore, alters the genome structure of their hosts (Bickmore and Sumner 1989; Kido et al. 1991). Changes in genome structure are prone to bring about rapid chromosomal rearrangements (Bernardi 1993; Hastie and Allshire 1989; Wichman et al. 1991). Since repetitive elements are easily detectable, they are good chromosomal and phylogenetic landmarks (Greene and Seeb 1997; Murata et al. 1993; Wichman et al. 1991).

Karyotypes of cyprinid fishes have been intensively studied, and most of them have karyotypes rich in bi-armed chromosomes. *Rhodeus atremius atremius*, *R. a. fangi* and *R. a. suigensis*, however, have a quite unusual karyotype, with many mono-armed chromosomes (Ojima

et al. 1973; Ueda et al. 1996). Congeneric *R. sinensis* (= *R. lighti* and *R. uyekii*) and *R. ocellatus*, on the other hand, have typical karyotype rich in bi-armed chromosomes (Ojima et al. 1973; Ueda et al. 1997). Isolation of chromosomal landmarks will facilitate further studies on structural changes of chromosomes related to such unusual karyotype divergence. For this purpose we tried to isolate repetitive sequences from these three *Rhodeus* bitterlings. The isolated sequences are similar to each other and may be the products of relatively recent amplification events.

MATERIALS AND METHODS

We tried to isolate repetitive sequences from a male of *Rhodeus atremius suigensis* from the Yalu River, a female of *R. sinensis* from the Nen River, and a male of *Rhodeus ocellatus ocellatus* from Putian, China (Table 1). Extraction of genomic DNAs from muscle tissue followed Asahida et al. (1996). We also used several bitterlings from China, Korea and Japan for comparative materials.

Complete *Hind*III-digests of genomic DNAs of *R. a. suigensis*, *R. sinensis*, and *R. o. ocellatus* were ligated with the compatible site of cloning vector pBluescriptIIS(+)(Stratagene, La Jolla, CA, USA) and cloned in DH5 α FIQ (Life Technologies, Rockville, MD, USA). Screening of

Edited by Toshihiko Shiroishi

* Corresponding author. E-mail: ksaitoh@myg.affrc.go.jp

Table 1. Fish used in this study. Localities are arranged from southern China to Japan within each species.

Species	Abbreviation	Locality	Sex
<i>Rhodeus atremius fangi</i>	<i>Rafa</i>	Min River, Fujian, China	Unknown
<i>R. a. suigensis</i>	<i>Rasu</i>	Yalu River, Liaoning, China	Male *
		Tal River, Jungju, Jungchongbuk-do, Korea	Female
		Namhan River, Janghowon, Gyonggi-do, Korea	Unknown
			Male
			Female
<i>R. a. atremius</i>	<i>Raat</i>	Baba River, Saga, Japan	Male
			Female
<i>R. sinensis</i>	<i>Rsi</i>	Small stream in Putian, Fujian, China	Female
		Small stream in Shanghai, China	Female
		Nen River, Heilongjiang, China	Two females **
		Ansong River, Gyonggi-do, Korea	Male
			Female
<i>R. ocellatus ocellatus</i>	<i>Rooc</i>	Small stream in Putian, Fujian, China	Male *
		Sungari River, Harbin, Heilongjiang, China	Male
			Female
		Tuwol River, Juksan-myon, Gyonggi-do, Korea	Male
			Female
<i>R. o. kurumeus</i>	<i>Roku</i>	Baba River, Saga, Japan	Male
			Female
		Small pond in Osaka, Japan	Male
			Female
<i>R. sericeus</i>	<i>Rse</i>	Nen River, Heilongjiang, China	Male
			Female
<i>Tanakia limbata</i>	<i>Tli</i>	Yabe River, Fukuoka, Japan	Female
<i>Acheilognathus rhombeus</i>	<i>Arh</i>	Floodplain pool in Miyagi, Japan	Female
<i>A. macropterus</i>	<i>Ama</i>	Nen River, Heilongjiang, China	Female

* Individuals from which we isolated repetitive sequences.
** We tried to isolate repetitive sequence from one of them, but the attempt was unsuccessful.

Fig. 1. Map (upper panel) and predicted secondary structure of the consensus sequence (lower panel) of the repetitive elements isolated from Chinese bitterlings. Numbers at the right margin of the map indicate lengths of clones in bp. Regions (a–c) show homology with sequences in some other vertebrates (see Fig. 2). Broad lines on the secondary structure diagram indicate positions of palindromes over 8 nucleotides long. Overlapping palindromes (underlines in inset) indicate the possibility of alternative foldings. The consensus sequence and alignment are available upon request.

repetitive DNAs followed Naruse et al. (1992). Prior to labeling the plasmid clone, PCR amplification of the insert sequence with KS and SK primers was employed to cut out the vector sequence. All the labeling, hybridization and detection techniques followed the supplier's recommendations (dig-system; Boehringer-Mannheim, Germany). The post-hybridization washing conditions were 2×SSC, 0.1% SDS (1×SSC; 0.15M NaCl, 0.015M Na₃-citrate) for 10min at room temperature and 0.1×SSC, 0.1% SDS for 5min twice at 65°C. To estimate signal intensity, membranes were scanned with IX-Scan ver. 3.5 on a model IX4025 with the linear sensitivity setting (Canon Sales, Tokyo, Japan) and quantified using NIH Image ver. 1.61 (National Institute of Health, Bethesda, MD, USA). Isolated clones were sequenced from both strands, including one or two deletion subclones using a dye-terminator cycle sequencing kit (Amersham-Pharmacia, Bucks, England) and an ABI373 automated sequencer (Perkin-Elmer, Norwalk, CT, USA).

We have deposited sequence data obtained in this study in the DDBJ/EMBL/GenBank database under the following accession numbers: AB036406–AB036413.

RESULTS

We first isolated 21 clones from the *Rhodeus atremius suigensis* individual which showed strong signals upon probing with total genomic DNA of that individual. Plasmid DNAs extracted from these clones were Southern blotted and probed with one of these clones (Rs20) which showed the strongest signal upon the colony hybridization. Two other clones (Rs12 and Rs15) also showed strong hybridization signals with that clone. We then decided to carry out further characterization of a possible highly repetitive DNA family represented by these clones. Two more clones (Rs23 and Rs24) were isolated by colony hybridization with Rs20 as a probe.

We also isolated the homologous elements from conge-

		Match%	Over
a AF045382 <i>Catla catla</i> microsatellite CcatG2 sequence		78.9	114
Rs.consensus	TAGCAA-CCAAATGCAAAGCGAG-ATATCTTCAGTTCTGAATGTCGTAGAGGCGTTGTGG	847	
AF045382	C.....CC..NCA·AN·C.....AA.....AC.....A·A·G·G·	408	
Rs.consensus	TTGGTTTATATCGTTCATACTGACAAGCAGCCTTTGAAGTGACATCAATGGCAG	901	
AF045382	N·A.....G.....GT··C.....T.....	462	
b AF154992 Zebrafish kidney cDNA random primed, RZPD #576		76.1	159
Rs.consensus	AAAATC-CCATTGACTTAACATTGCGACGAACTTTGACGATTAATAACTCGGAGCGAGGA	382	
AF154992	··G··T······T····G·A·C······A·C·C·G·G·C···N·T·A·	294	
Rs.consensus	ATTTGTACAAAGATGAACTCGCCACATTTGAAGAGGGTGACAGCCTCTGTAAGAACATA	442	
AF154992	T·CA·G··T··GG·T···A···C··T·G···TG····G·A·G·····	354	
Rs.consensus	CCACACAATGGGGTAAAAGTTGCACCCCTGGGGTGCTAG	481	
AF154992	··C·GA·A····T····T····	393	
c U67076 <i>Xenopus laevis</i> zinc finger protein XFDL156 gene		72.6	219
Rs.consensus	CAGGGGCAAAAAAGCAGCCGAAAAATCCCATTGACTTAACATTGCGACGAAC-TTTGACG	361	
U67076<	·G····TT··C··C··C··G····A··-····A·G·C··T····	1440	
Rs.consensus	ATTAATAACTCGGAGCGAGGAATTTGTACAAA-GATGAACTCGCCACATTTGAAGAGGG	420	
U67076<	GA·TC··G·G·A··A·G·A·TC····G··C·T·A··T·TA······T	1380	
Rs.consensus	TGACAGCCTCTGTAAGAACATACCACACAATGGGGTAAAAGTTGCACCCCTGGGGTGCTA	480	
U67076<	·TG·GA··G··C··TG····C····CGA····T····TT····-····CAGG·	1321	
Rs.consensus	GAGCTGCCCAAATTTGCCCATAGAGAT-GAATGGGGAA	518	
U67076<	··G·C········T··CT·AT····	1282	

Fig. 2. Regions which are highly homologous with other vertebrate sequences. <: complementary sequence.

neric *R. sinensis* and *R. ocellatus ocellatus*. From genomic DNA of *R. o. ocellatus* we isolated three clones (Ro2, Ro3 and Ro5) showing strong hybridization signals with Rs20. For *R. sinensis*, however, isolation of homologous elements was unsuccessful. No strong or moderate signal was detectable upon colony hybridization with a clone isolated from *R. a. suigensis*.

Sequencing showed that the structures of the eight clones were close to each other except for frequent indels (Fig. 1, upper panel). Seven clones out of eight contained a {TTA}_n (n=2 to 11) trinucleotide repeat near one end. A 220bp indel was recognizable near the opposite end. The clones showed 95.0 to 99.8% identity with each other, excluding indels and the internal repeat. The apparent transition/transversion ratio was 3.00.

There are several short and a few longer palindromic sequences as well as distantly separated inverted repeat regions. Rigid secondary structures are accordingly predictable for single-stranded sequences (Fig. 1, lower panel) ($\Delta G=-315.0$ to -455.9 kcal/mol for normal and -322.4 to

-464.3 kcal/mol for complementary sequences). As for the consensus sequence, the predicted values are -512.2 and -513.3 kcal/mol, respectively. Overlaps between palindromes or between palindromes and inverted repeat regions may indicate the capability of alternative modes of folding (Fig. 1, inset).

Two regions show similarity with other known vertebrate sequences (Fig. 2). The region from 790 to 901 in the consensus sequence is similar to the Indian major carp (*Catla catla*) microsatellite-flanking region (Naish and Skibinski 1998). The region around 300 to 500 is similar to a cDNA of unknown function from zebrafish (*Danio rerio*) (Look and Holloway, unpublished; acc. no. AF154992) and to *Xenopus laevis* zinc finger protein cDNA (Bellefroid et al. 1997). The zebrafish sequence that is similar to the bitterling repeat contains many stop codons, indicating that this is an untranslated region. The portion of the *Xenopus* sequence showing similarity with the bitterling repeat is in the upstream untranslated region. There are few palindromic sequences within

Fig. 3. Southern hybridization patterns of highly repetitive elements from several bitterlings. Left panel shows partial *Hind*III (0.1 to 1U for 1h) digestion of *Rhodeus atremius suigensis* individual from which we isolated the element. λ : molecular weight markers (*Eco*T14I-digested λ -DNA). Samples in the right panel are complete *Hind*III-digested (15U for 1.5h) genomic DNAs of bitterlings from several localities. Each lane contains about 1 μ g of genomic DNA. Numbers in the left margin indicate estimated lengths of bands. See Table 1 for abbreviations.

Fig. 4. Dot blot hybridization indicating relative copy number of the repetitive elements. Genomic DNAs are of individuals from which we tried to isolate the elements. For genomic DNA and λ -DNA (negative control), 1 to 100ng of DNA was spotted. For the plasmid (Rs24), 0.03 to 30ng was spotted. See Table 1 for abbreviations.

these regions homologous with other known sequences, but some complementarities are present in their flanking regions. The existence of palindromes and internal complementarities is characteristic of the MITE (miniature inverted-repeat transposable element) family (Izsvák et al. 1999), but the repeat of our study is much longer.

Genomic Southern blotting with labeled Rs24 (the longest clone) as a probe revealed two major elements of about 1.0 and 1.3kb (Fig. 3). The 220bp indel among the clones corresponds approximately with the two size groups detected by Southern blotting. Smeared hybridization signals of moderate strength indicate the presence of interspersed elements homologous with the probe. A partial digestion experiment revealed a very weak ladder pattern, indicating that some portion of the elements constitute a tandem array (Fig. 3, left panel). A band of about 2.3kb, a putative heterodimer of the 1.0 and 1.3kb elements, was barely visible along with a 2.1kb band, a possible homodimer of 1.0kb elements. Bands at positions expected for other oligomers were also weakly visible. While there was no indication of variabilities within *R. atremius*, *R. ocellatus* showed variation within the species, i.e., a lack of a strong signal of the 1.3kb band in *R. o. kurumeus*. *R. sinensis* exhibited weak smeared and 1.0kb signals. *R. sericeus* and bitterlings of other genera exhibited only weak smeared signals.

The amount of the repetitive elements in bitterling genomes estimated by dot blot hybridization with Rs24 probe

ranged from 0.29 to 1.36% of the entire genomes: $0.81 \pm 0.015\%$ for *R. a. suigensis*, $1.36 \pm 0.093\%$ for *R. o. ocellatus* and $0.29 \pm 0.023\%$ for *R. sinensis* (Fig. 4). The DNA content of *R. o. ocellatus* was determined to be 1.12pg/haploid (55.8% of goldfish; Ojima et al. 1972, Hinegardner and Rosen 1972). From this value and a postulated average length of the repeat unit being 1.2kb, the approximate copy number of this repeat in the *R. o. ocellatus* haploid genome is 11000.

DISCUSSION

We have isolated a highly repetitive family of partially tandem and partially interspersed elements. The former component shows the periodical presence of *Hind*III recognition sites, and the latter has two internal *Hind*III sites. Sequencing and Southern blotting revealed two major length groups with several indels. Whether each cloned sequence belongs to tandem or interspersed components is unclear, but the consensus sequence may represent the family of elements regardless of their interspersion. The banding pattern on Southern blot membranes indicates that the repeat family is specific to a few species within the genus *Rhodeus* and that it was relatively recently amplified. The signal intensities are, however, different from species to species. The differences in the banding pattern between *R. ocellatus* and *R. o. kurumeus* may indicate divergence

within the species, as indicated by a few studies of mtDNA analysis (Miyake et al. in press; Kawamura et al. unpublished).

The repeat family shows some similarities with other fish and higher vertebrate sequences. Since similarity of these sequences along a stretch of over 100bp by chance is hardly likely, these similarities indicate homology. Therefore, incorporation of extant genomic sequences of the host itself into the repeat family happened during the evolutionary course of the family.

Rigid secondary structures inferred from palindromic and inverted repeat sequences might facilitate production of indels through replication errors (Izsvák et al. 1999; Levinson and Gutman 1987; Stephan 1986). Possible single-copy sequence-derived regions might transform into repetitive sequences with some complementarity with the flanking sequences.

Dot blotting revealed rapid changes in copy number relative to genome size among some *Rhodeus* species. *R. atremius* and *R. ocellatus* have similar copy number, but mtDNA phylogeny shows that *R. atremius* is closer to *R. sinensis* than to *R. ocellatus* (Okazaki et al. in press). If the mtDNA phylogeny is correct, we can suggest the following two alternative explanations. (1) The repetitive family of this study originated in a common ancestor of the three species in *Rhodeus* and rapidly amplified in *R. atremius* and *R. ocellatus* genomes independently. (2) Genome size increase occurred in the *R. atremius-sinensis* lineage and this tendency is especially conspicuous in *R. sinensis*. Regardless of whether either of these two alternatives is correct, rapid changes in genome structure occurred with diversification of *Rhodeus* bitterlings.

Since repetitive sequences are easily detectable, they are useful chromosomal and phylogenetic landmarks. This is especially important for bitterlings that show unusual patterns of chromosomal rearrangements consisting of frequent inversions from bi-armed to mono-armed chromosomes (Naoi et al. unpublished). The repetitive family isolated in this study may be a useful landmark for future chromosomal studies.

We thank Y. Xie, Research Institute for Freshwater Fisheries of Liaoning Province, Z. Weng, Ichthyological Division, Shanghai Fisheries University and H. Naoi, Department of Biology, Faculty of Education, Utsunomiya University for their help in fish collection in the field. This study was partly supported by the Grants-in-Aid from the Japan Ministry of Education, Science, Sports and Culture (B(2)10041156, B(1)12575009).

REFERENCES

- Asahida, T., Kobayashi, T., Saitoh, K., and Nakayama, I. (1996) Tissue preservation and total DNA extraction from fish stored at ambient temperature using buffers containing high concentration of urea. *Fish. Sci.* **62**, 727–730
- Bellefroid, E., Bourguignon, C., Bouwmeester, T., Rausch, O., Blumberg, B., and Pieler, T. (1997) Transcription regulation and alternative splicing of an early zygotic gene encoding two structurally distinct zinc finger proteins in *Xenopus laevis*. *Mech. Dev.* **63**, 99–108.
- Bernardi, G. (1993) Genome organization and species formation in vertebrates. *J. Mol. Evol.* **37**, 331–337
- Bickmore, W. A. and Sumner, A. T. (1989) Mammalian chromosome banding – an expression of genome organization. *Trend. Genet.* **5**, 144–148
- Britten, R. J. and Kohne, D. E. (1968) Repeated sequences in DNA. *Science* **161**, 529–540
- Doolittle, W. F. and Sapienza, C. (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**, 601–603
- Greene, B. A. and Seeb, J. E. (1997) SINE and transposon sequences generate high-resolution DNA fingerprints, “SINE prints,” that exhibit faithful Mendelian inheritance in pink salmon (*Oncorhynchus gorbuscha*). *Mol. Mar. Biol. Biotechnol.* **6**, 328–338
- Hastie, N. D. and Allshire, R. C. (1989) Human telomeres: fusion and interstitial sites. *Trend. Genet.* **5**, 326–331
- Hinegardner, R. and Rosen, D. E. (1972) Cellular DNA content and the evolution of teleostean fishes. *Am. Nat.* **106**, 621–644
- Hori, H., Suzuki, M., Inagaki, H., Ohima, T., and Koga, A. (1998) An active Ac-like transposable element in teleost fish. *J. Mar. Biotechnol.* **6**, 206–207
- Izsvák, Z., Ivics, Z., Garcia-Estefania, D., Fahrenkrug, S. C., and Hackett, P. B. (1996) DANA elements: A family of composite, tRNA-derived short interspersed DNA elements associated with mutational activities in zebrafish (*Danio rerio*). *Proc. Natl. Acad. Sci. USA* **93**, 1077–1081
- Izsvák, Z., Ivics, Z., Shimoda, N., Mohn, D., Okamoto, H., and Hackett, P. B. (1999) Short inverted-repeat transposable elements in teleost fish and implications for a mechanisms of their amplification. *J. Mol. Evol.* **48**, 13–21
- Kido, Y., Aono, M., Yamaki, T., Matsumoto, K., Murata, S., Saneyoshi, M., and Okada N. (1991) Shaping and reshaping of salmonid genomes by amplification of transfer RNA-derived retroposons during evolution. *Proc. Natl. Acad. Sci. USA* **88**, 2326–2330
- Levinson, G. and Gutman, G. A. (1987) Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**, 203–221
- Miyake, K., Tachida, H., Oshima, Y., Arai, R., Kimura, S., Imada, N., and Honjo, T. (2000) Genetic variation of the cytochrome *b* gene in the rosy bitterling, *Rhodeus ocellatus* (Cyprinidae) in Japan. *Ichthyol. Res.* (in press)
- Murata, S., Takasaki, N., Saitoh, M., and Okada, N. (1993) Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. *Proc. Natl. Acad. Sci. USA* **90**, 6995–6999
- Naish, K. A. and Skibinski, D. O. F. (1998) Tetranucleotide microsatellite loci for Indian major carp. *J. Fish Biol.* **53**, 886–889
- Naruse, K., Mitani, H., and Shima, A. (1992) A highly repetitive interspersed sequence isolated from genomic DNA of the medaka, *Oryzias latipes*, is conserved in three other related species within the genus *Oryzias*. *J. Exp. Zool.* **262**, 81–86
- Ojima, Y., Hayashi, M., and Ueno, K. (1972) Cytogenetic studies in lower vertebrates. X. Karyotype and DNA studies in 15 species of Japanese Cyprinidae. *Jpn. J. Genet.* **47**, 431–440
- Ojima, Y., Ueno, K., and Hayashi, M. (1973) Karyotypes of the acheilognathine fishes (Cyprinidae) of Japan with a discussion of phylogenetic problems. *Zool. Mag.* **82**, 171–177
- Okazaki, M., Naruse, K., Shima, A., and Arai R. (2000) Phylogenetic relationships of bitterlings based on mitochondrial 12S

- ribosomal DNA sequences. J. Fish Biol. (in press)
- Orgel, L. E. and Crick, F. H. C. (1980) Selfish DNA: the ultimate parasite. *Nature* **284**, 604–697
- Stephan, W. (1986) Recombination and the evolution of satellite DNA. *Genet. Res.* **47**, 167–174
- Ueda, T., Arai, R., Akai, Y., Ishinabe, T., and Wu, H. (1996) Karyological study of a bitterling *Rhodeus atremius* (Teleostei: Cyprinidae) collected in Zhejiang, China. *Cytobios* **86**, 265–268
- Ueda, T., Mashiko, N., Takizawa, H., Akai, Y., Ishinabe, T., Arai, R., and Wu, H. (1997) Ag-NOR variation in chromosomes of Chinese bitterlings, *Rhodeus lighti* and *Tanakia himantegus* (Cypriniformes, Cyprinidae). *Ichthyol. Res.* **44**, 302–305
- Wichman, H. A., Payne, C. T., Ryder, O. A., Hamilton, M. J., Maltbie, M., and Baker, R. J. (1991) Genomic distribution of heterochromatic sequences in equids: Implications to rapid chromosomal evolution. *J. Hered.* **82**, 369–377