Transposable Elements in Medaka Fish

Authors: Koga, Akihiko, Sakaizumi, Mitsuru, and Hori, Hiroshi

Source: Zoological Science, 19(1) : 1-6

Published By: Zoological Society of Japan

URL: https://doi.org/10.2108/zsj.19.1

ZOOLOGICAL SCIENCE **19**: 1–6 (2002)

[REVIEW]

Transposable Elements in Medaka Fish

Akihiko Koga[1*], Mitsuru Sakaizumi[2] and Hiroshi Hori[1]

[1]*Division of Biological Sciences, Graduate School of Science, Nagoya University, Nagoya 464-8602, Japan*
[2]*Department of Environmental Science, Faculty of Science, Niigata University, Niigata 950-2181, Japan*

ABSTRACT—DNA-based transposable elements appear to have been nearly or completely inactivated in vertebrates. Therefore the elements of the medaka fish *Oryzias latipes* that still have transposition activity provide precious materials for studying transposition mechanisms, as well as the evolution, of transposable elements in vertebrates. Fortunately, the medaka fish has a strong background for genetic and evolutionary studies. The advantages of this host species and their elements, together with results so far obtained, are here described.

Key words: transposable elements, medaka fish, evolution, biotechnology

DNA-based transposable elements

Four major classes of transposable elements are known: viral family retrotransposons, long interspersed nuclear elements (LINEs), short interspersed nuclear elements (SINEs) and DNA-based transposable elements. The last class, also called terminal-inverted-repeat elements, is different from the others in that elements move directly from DNA to DNA, that is, without forming RNA intermediates. The enzyme that is central or solely required for their transposition reaction is called a transposase and is considered to act mainly on their terminal and/or subterminal regions. Transposable elements are present in multicopies in host genomes and, in many cases, albeit not all, autonomous copies exist including a gene for their own transposase. Copies that do not carry the transposase gene or carry an imperfect gene are called nonautonomous copies, these being thought to be produced from autonomous copies by errors during transposition or subsequent mutational sequence changes. As long as the transposase cannot distinguish autonomous from nonautonomous copies, which so far appears to be true, the evolutionary scenario they would be expected to follow is first increase in frequency of nonautonomous elements, subsequent loss of transposition activity, and ultimate extinction from the host genome (cf. Lohe *et al.,* 1995; Hartl *et al.,* 1997).

Several DNA-based transposable elements have been identified in vertebrates, analogous to many other organisms. However, autonomous copies are exceedingly rare, which suggests the situation where most vertebrate DNA-based elements are in the final or nearly the final stage of the scenario described above. Inferences from the results of the human genome sequencing project are consistent with this view (cf. International Human Genome Sequencing Consortium 2001). For studying transposition mechanisms, as well as evolution, the few elements in teleost fishes demonstrated or suggested to be still active offer unique materials. The authors have been working on identification and examination of active elements of the medaka fish *O. latipes*. The present review describes results so far obtained on various aspects of medaka fish transposable elements, such as structure, transposition mechanisms, evolution and application to biotechnology.

Medaka fish

The medaka fish is a freshwater teleost native to Eastern Asia, including China, Japan and Korea. It is, together with zebrafish, a superior model animal for vertebrate genetics and developmental biology because of its small genome size (800 Mb), small body size (4 cm), short generation time (3 months), high reproductive rate (10 to 20 eggs per female every day) and transparent egg chorion (Iwamatsu 1997; Ishikawa 2000; Packer 2001; also see Medakafish Home-

* Corresponding author: Tel. +81-52-789-2506;
FAX. +81-52-789-2974.
E-mail: koga@bio.nagoya-u.ac.jp

page at http://biol1.bio.nagoya-u.ac.jp:8000/). In addition, a strain having completely transparent bodies has recently been generated by accumulating deficient alleles at several pigmentation loci (Wakamatsu *et al.,* 2001). Another remarkable advantage is the in-depth knowledge of closely related fishes. Fourteen species are so far identified in the genus Oryzias and their detailed phylogenetic relationship is known (Naruse *et al.,* 1993; Naruse, 1996; Koga *et al.,* 2000). This provides a strong basis for evolutionary studies comparative to that in the *melanogaster* species subgroup of Drosophila. A great amount of genetic variation is known to exist not only among but also within species (Sakaizumi *et al.,* 1983) and many inbred strains reflecting geographic differences have been prepared (Hyodo-Taguchi and Sakaizumi, 1993). In relation to identifying active transposable elements, a useful infrastructure is provided by the collection of various body color mutant lines available mainly due to the efforts of Hideo Tomita of Nagoya University. His professional experience and patience have allowed identification of various genetic variants of the medaka fish and establishment of mutant lines. These are all spontaneous mutations and still maintained in Nagoya and at other institutions (Tomita, 1975, 1992).

The *Tol1* element

There are several different albino mutants in Tomita's collection. He was convinced (personal communication) that some of them are due to defective tyrosinase genes. Tyrosinase is the key enzyme in melanin biosynthesis. Because structural analysis of spontaneous mutant genes has long been the main procedure for identifying new transposable elements, we first cloned the wild-type tyrosinase gene of the medaka fish (Inagaki *et al.,* 1994) and then examined its structure in candidate albino strains. The *Tol1* element (transposable element of *Oryzias latipes*, 1; GenBank accession number D42062) was found as a result of this analysis. It had been inserted in the first exon of the tyrosinase gene, making the gene nonfunctional (Koga *et al.*, 1995).

The particular *Tol1* copy found in the tyrosinase gene was denoted *Tol1-tyr*. It is 1.9 kb in length, carries terminal inverted repeats of 14 bp, and is flanked by an 8-bp target site duplication (Fig. 1.). No meaningful open reading frame (ORF) is evident in *Tol1-tyr*. Mainly for this reason, it is not clear which other transposable elements *Tol1* is related to. We have estimated the copy number of *Tol1* to be about 80 per diploid genome (Koga *et al.,* 1995).

While the human tyrosinase gene is more than 70 kb in length (Ponnazhagan *et al.,* 1994), that in the medaka fish is as small as 5 kb, primarily because of the smaller sizes of introns (Inagaki *et al.,* 1998). The entire defective gene of the albino strain, containing the 1.9-kb *Tol1* copy, is within the carrying capacity of a lambda phage vector. We obtained such a clone and replaced a small part of the first exon, in which *Tol1* is included, with the corresponding region of the wild-type gene. Introduction of this modified clone into fertilized eggs of the albino strain resulted in melanin pigmentation in the skin of hatched fish. There was no nucleotide sequence difference in the manipulated region between the wild-type and the defective genes, except for the *Tol1* insertion and its flanking target site duplication (unpublished). These results indicate that the particular *Tol1* insertion event in the tyrosinase gene is not so old as to lead to accumulation of other mutations on the already defective tyrosinase gene. In other words, *Tol1* is, or was until recently, active with respect to transposition.

More direct evidence for its transposition activity has recently been obtained. We exposed embryos of the albino strain to UV, and conducted PCR with a pair of primers encompassing the *Tol1* insertion point. The length of a major PCR product was longer by 1.9 kb than the product with a wild-type fish. However, in some cases, minor products of a length identical, on agarose gels, to the wild-type form were observed. Cloning and sequencing of these products revealed excision footprints, remnants of *Tol1* terminal regions or loss of nucleotides originally belonging to the

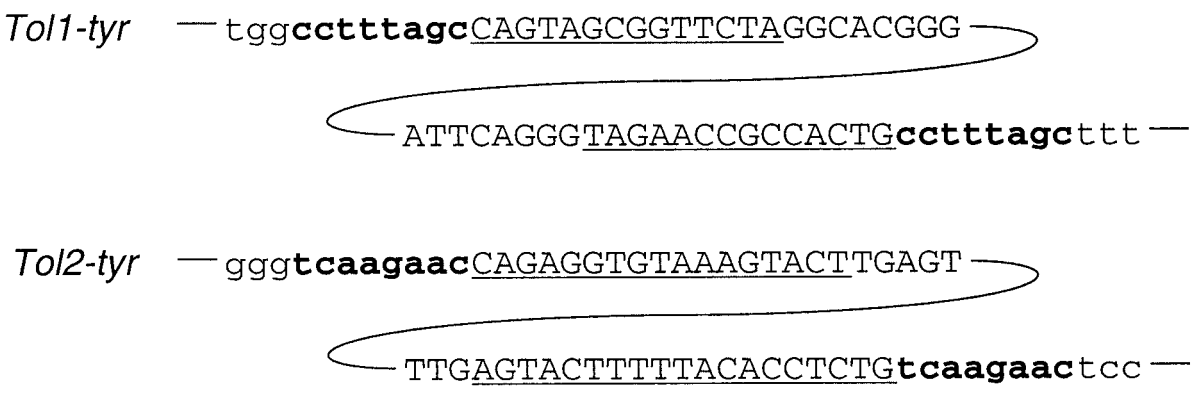

Fig. 1. Nucleotide sequences of *Tol1* and *Tol2* terminal regions. *Tol1-tyr* and *Tol2-tyr* are specific copies of *Tol1* and *Tol2*, respectively, inserted in the tyrosinase gene. Nucleotide sequences of their terminal regions are shown in uppercase letters. *Tol1* has terminal inverted repeats of 14 bp each and *Tol2* carris terminal inverted repeats of 17 bp and 19 bp, indicated by underlining. The sequences of the tyrosinase gene are shown in lowercase letters. Both elements are flanked by 8-bp target site duplications indicated by boldface letters.

tyrosinase gene (unpublished). Such footprints are observed with many DNA-based transposable elements (cf. Pohlman *et al.,* 1984; Bryan *et al.,* 1990).

From the above results, an autonomous copy that is presumed to carry a transposase gene would be expected to be present somewhere in the genome. However, no such copy has been detected among more than 20 copies we have so far cloned and sequenced.

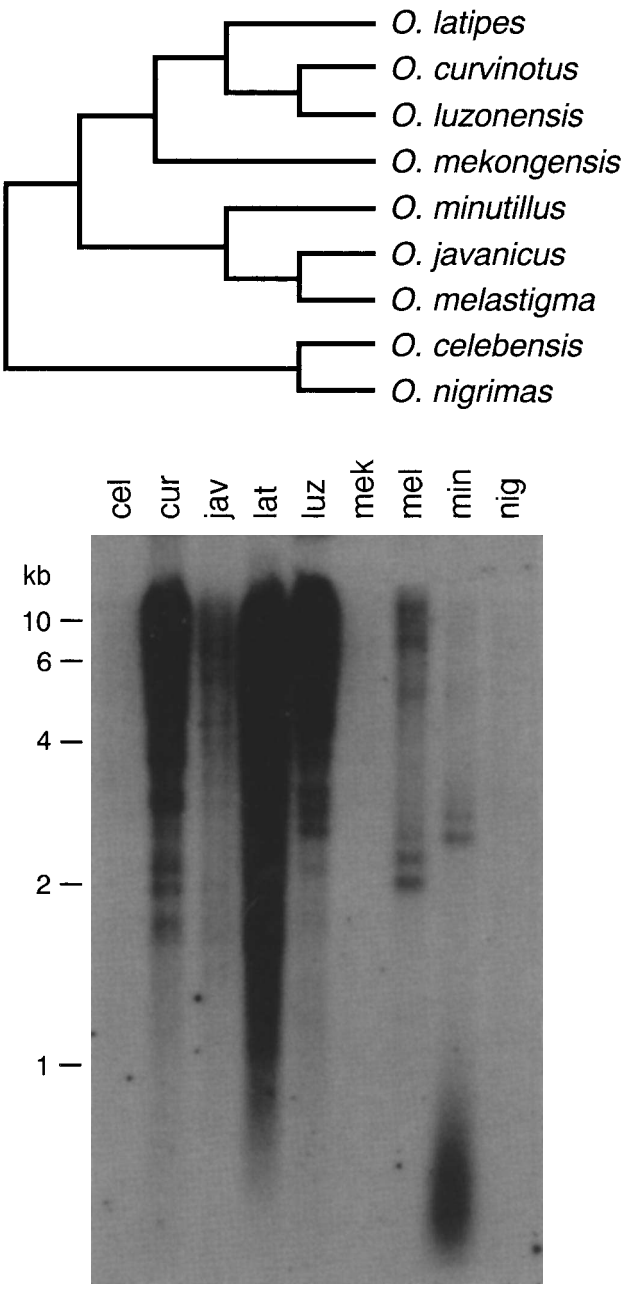

Fig. 2. Distribution of *Tol1* among species. Phylogenetic trees with nucleotide sequence data were constructed by Naruse *et al.* (1993) and Naruse (1996). The scheme was redrawn from these references for the nine species examined. Genomic Southern blot hybridization to examine the distribution of *Tol1* among species was conducted as described (Koga *et al.,* 1995), using the entire *Tol1-tyr* element as probe. The species names are abbreviated with their first three letters. Hybridization signals are not evident in the the lane for *O. mekongensis* on this autodiagram, but longer exposure revealed about ten bands. Along the left margin are indicated the sizes and mobilities of the size marker DNA fragments.

Tol1 occurs in seven of the nine Oryzias species we have examined by genomic Southern blot analysis (Fig. 2). The probe used was the 1.9-kb entire *Tol1-tyr* copy cloned from the medaka fish *O. latipes*. Hybridization signals became weaker with the phylogenetic distance from *O. latipes*, with the exception of *O. mekongensis*. This result suggests that *Tol1* was present in the common ancestor of the Oryzias species. We have also conducted more analytical Southern blot experiments, in which different parts of the 1.9-kb *Tol1* copy were used separately as probes. We observed several species-specific structures. For example, a central 0.4-kb region of *Tol1-tyr* hybridized to virtually all *Tol1* copies in *O. latipes* but did not produce any hybridization signals with *O. javanicus*, indicating that all the *Tol1* copies of the latter species lack this portion (unpublished). These results suggest that, although *Tol1* appears to be an ancient resident in the genus Oryzias, its transposition occurred after species divergence.

The *Tol2* element

Tol2 was identified by the same approach, as adopted for *Tol1*, in a different albino strain, that is, a different tyrosinase allele (Koga *et al.,* 1996; GenBank D84375). It had been inserted in the last, fifth exon of the tyrosinase gene. It is 4.7 kb in length, carries terminal inverted repeats of 17 bp and 19 bp, and is flanked by an 8-bp target site duplication (Fig. 1). This element belongs to the *hAT* transposable element family (Calvi *et al.,* 1991; Atkinson *et al.,* 1993) that includes *hobo* of Drosophila, *Activator* of maize and *Tam3* of snapdragon. *Tol2* was directly demonstrated to be active soon after its discovery. We carried out PCR with primers encompassing the *Tol2* insertion point, and PCR products with the same length as the product from the wild-type fish were readily obtained without any treatment such as exposure to UV. In these PCR products, the element was found to have been removed and footprints were observed.

Tol2 has four ORFs that roughly correspond to four exons making up the mRNA for the *Tol2* transposase (Koga *et al*., 1999; Koga and Hori, 2000), which catalyzes both excision and insertion of *Tol2*. Microinjection of fish eggs with *in vitro* synthesized mRNA for the transposase triggers transposition of a co-injected *Tol2* clone (Koga and Hori, 2000; Kawakami *et al.,* 2000). This opened a way to the development of a gene transfer vector. We embedded the green fluorescent protein (GFP) gene in *Tol2* and injected this clone into fertilized eggs together with the transposase mRNA. The GFP gene was successfully introduced into the fish genome and transmitted to subsequent generations. Sequencing analysis demonstrated the entire *Tol2* element, including the GFP gene, to have been integrated into the chromosomes through its transposition (Koga *et al.,* 2001).

The *Tol2* clone, modified to carry the GFP gene and other necessary components, was 9.0 kb in length. Therefore, *Tol2* still retains transposition activity even when it is as large as 9.0 kb. Larger versions are now being tested to determine the limitation in size.

Regarding variation among copies in the genome and the distribution among species, *Tol2* presents a striking contrast to *Tol1*. We examined the presence/absence of *Tol2* in ten Oryzias species and found it only in two, the medaka fish *O. latipes* and the Hainan medaka fish *O. curvinotus*, the latter being an inhabitant of Southern China. These two species are relatively closely related to each other but nucleotide sequence variations exist on their chromosomes. Despite this variation, *Tol2* copies exhibit virtually no nucleotide sequence differences between the species. From results of quantitative analysis, we have proposed that *Tol2* was recently introduced into one of these two species from the other or into both of them from an unknown source (Koga *et al.,* 2000).

An LINE-like element

Nonautonomous copies of DNA-based elements are frequently shorter than autonomous copies due to deletion of internal regions. Therefore, a longer copy, if found, can be taken as a candidate for an autonomous copy. We therefore examined length variation of *Tol1* copies randomly collected from genomic libraries of *O. latipes* and *O. curvinotus*. Although no candidate has yet been obtained, a by-product of this search was an LINE-like element inserted in *Tol1*. One of the *Tol1* copies we examined was disrupted by an extra fragment of about 5 kb and this fragment proved to contain ORFs. The longest frame was for 388 amino acids and their sequence exhibited a similarity to that of the

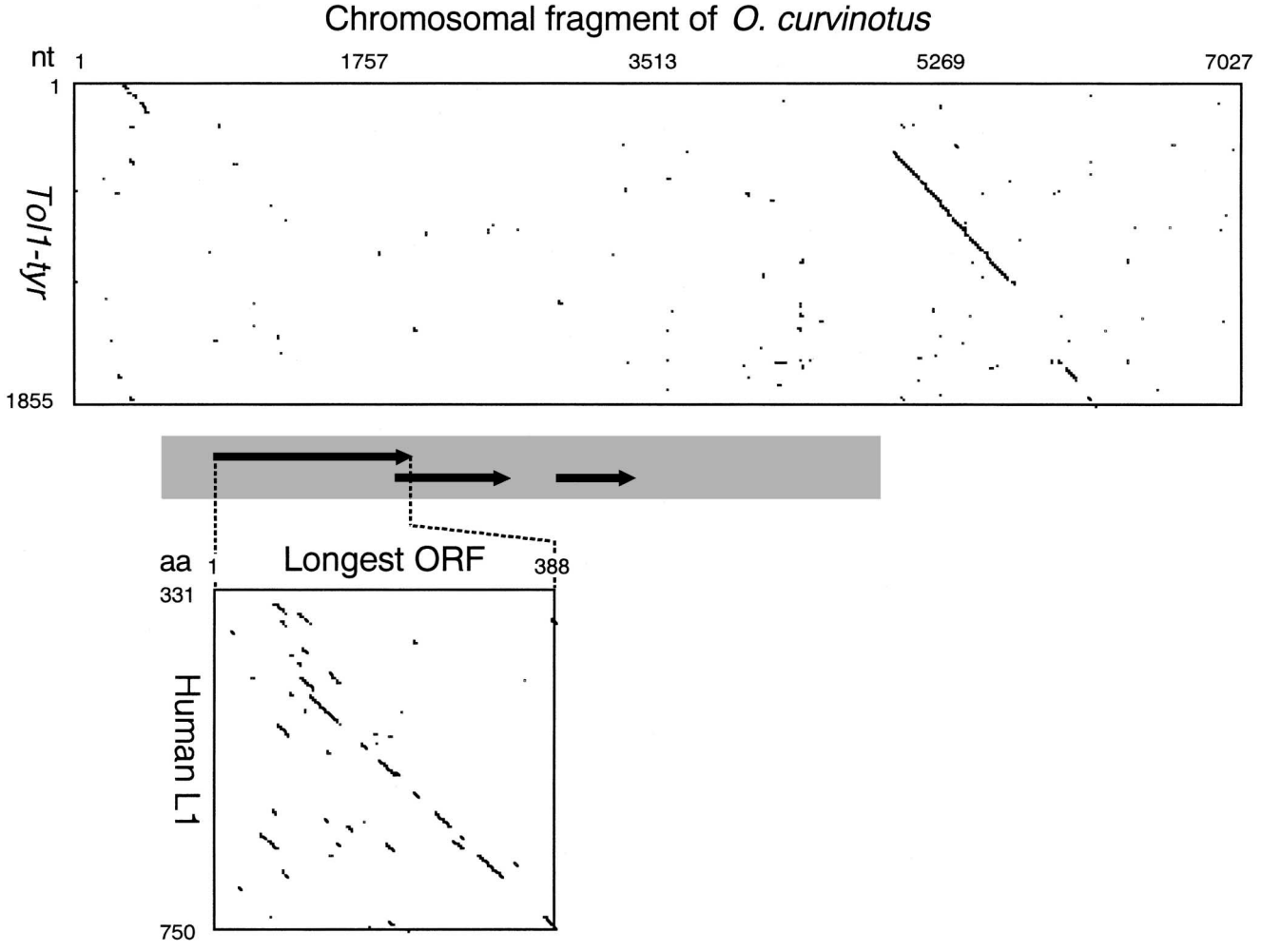

Fig. 3. The newly found LINE-like element. The upper dot matrix is for nucleotide sequences (nt) between the genomic fragment containing a *Tol1* copy of *O. curvinotus* (abscissa) and the first found *Tol1* copy of *O. latipes*, *Tol1-tyr* (ordinate). The criterium for matching was a 80% match over a window of 20 nucleotides. The *Tol1* copy of *O. curvinotus* contains an insertion sequence, which is represented by the shaded box. The arrows in the box are included ORFs. The lower dot matrix is for amino acid sequences (aa) between the longest ORF (abscissa) and part of the second ORF of the human L1 element (ordinate). The criterium for matching was a 30% match over a window of 20 amino acids. The human L1 element is quoted from GenBank X52234, and the part used here corresponds to its reverse transcriptase gene region.

reverse transcriptase gene of the human L1 (LINE-1) element (Fig. 3). This insertion sequence is present in multiple copies in the genomes of *O. latipes* and *O. curvinotus*. Because the ORF is not sufficiently long, the sequence is not likely to be a complete copy of LINE. An LINE element called *Swimmer 1* has already been identified in the medaka fish genome (Duvernell and Turner, 1998). The newly found LINE-like element, however, seems to differ from *Swimmer 1*.

Future studies

Evidently *Tol1* and *Tol2* are active with respect to transposition. Therefore, these elements can be expected to provide useful genetic tools applicable to vertebrates. In particular, *Tol2* belongs to the *hAT* transposable element family, members of which are widely distributed among organisms, including vertebrates (Esposito *et al.,* 1999; Rubin *et al.,* 2001). We have already demonstrated that *Tol2* works as a gene transfer vector in fish (Koga *et al.,* 2001). We are now trying to develop other types of genetic tools, such as an insertional mutagenesis agent and a cloning tag, first in fish and subsequently in other vertebrates.

The elements are also useful in studying the evolution of transposable elements because *Tol2* seems to be in a very early stage and *Tol1* in a relatively late stage of the probable scenario for transposable elements, and because a strong background for evolutionary studies has already been generated for the medaka fish.

ACKNOWLEDGMENTS

We owe much to Hideo Tomita of Nagoya University, who passed away in 1998, for his support of our work on transposable elements. Fish samples of various species were provided by H. Hashikawa, M. Sato and S. Susaki of the World Medaka Aquarium at Nagoya City Higashiyama Zoological Garden. This work was supported by grants no. 10216205 to A. K. and no. 11P0201 to H. H. from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

Atkinson PW, Warren WD, O'Brochta DA (1993) The *hobo* transposable element of Drosophila can be cross-mobilized in houseflies and excises like the *Ac* element of maize. Proc Natl Acad Sci USA 90: 9693–9697

Bryan G, Garza D, Hartl DL (1990) Insertion and excision of the transposable element mariner in Drosophila. Genetics 125: 103–114

Calvi BR, Hong TJ, Findley SD Gelbert WM (1991) Evidence for a common evolutionary origin of inverted repeat transposons in Drosophila and plants: hobo, Activator, and Tam3. Cell 66: 465–471

Duvernell DD, Turner BJ. (1998) *Swimmer 1*, a new low-copy-number LINE family in teleost genomes with sequence similarity to mammalian *L1*. Mol Biol Evol 15: 1791–1793

Esposito T, Gianfrancesco F, Ciccodicola A, Montanini L, Mumm S, D'Urso M, Forabosco A (1999) A novel pseudoautosomal human gene encodes a putative protein similar to *Ac*-like transposases. Hum Mol Genet 8: 61–67

Hartl DL, Lohe AR, Lozovskaya ER (1997) Modern thoughts on an ancyent marinere: function, evolution, regulation. Annu Rev Genet 31: 337–358

Hyodo-Taguchi Y, Sakaizumi M (1993) List of inbred strains of the medaka, *Oryzias latipes*, maintained in the Division of Biology, National Institute of Radiological Sciences. Fish Biol J Medaka 5: 29–30

Inagaki H, Bessho Y, Koga A, Hori H (1994) Expression of the tyrosinase-encoding gene in a colorless melanophore mutant of the medaka fish, *Oryzias latipes*. Gene 150: 319–324

Inagaki H, Koga A, Bessho Y, Hori H (1998) The tyrosinase gene from medakafish: transgenic expression rescues albino mutation. Pigment Cell Res 11: 283–290

International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921

Ishikawa Y (2000) Medakafish as a model system for vertebrate developmental genetics. Bioessays 22: 487–495

Iwamatsu T (1997) The Integrated Book for the Biology of the Medaka (in Japanese). Daigaku Kyoiku Publ., Okayama, Japan

Kawakami K, Shima A, Kawakami N (2000) Identification of a functional transposase of the *Tol2* element, an *Ac*-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. Proc Natl Acad Sci 97:11403–11408

Koga A. Hori H (2000) Detection of *de novo* insertion of the medaka fish transposable element *Tol2*. Genetics 156: 1243–1247

Koga A, Hori H, Sakaizumi M (2001) Gene transfer and cloning of flanking chromosomal regions using the medaka fish *Tol2* transposable element. Marine Biotech: in press

Koga A, Inagaki H, Bessho Y, Hori H (1995) Insertion of a novel transposable element in the tyrosinase gene is responsible for an albino mutation in the medaka fish, *Oryzias latipes*. Mol Gen Genet 249: 400–405

Koga A, Shimada A, Shima A, Sakaizumi M, Tachida H, Hori H (2000) Evidence for recent invasion of the medaka fish genome by the *Tol2* transposable element. Genetics 155: 273–281

Koga A, Suzuki M, Inagaki H, Bessho Y, Hori H (1996) Transposable element in fish. Nature 383: 30

Koga A, Suzuki M, Maruyama Y, Tsutsumi M, Hori H (1999) Amino acid sequence of a putative transposase protein of the medaka fish transposable element *Tol2* deduced from mRNA nucleotide sequences. FEBS Lett: 461: 295–298

Lohe AR, Moriyama EN, Lidholm DA, Hartl DL (1995) Horizontal transmission, vertical inactivation, and stochastic loss of *mariner*-like transposable elements. Mol Biol Evol 12: 62–72

Naruse K (1996) Classification and phylogeny of fishes of the genus Oryzias. Fish Biol J Medaka 8**:** 1–10

Naruse K, Shima A, Natsuda M, Sakaizumi M, Iwamatsu T (1993) Distribution and phylogeny of rice fish and their relatives belonging to the suborder Arianichthyoidei in Sulawesi, Indonesia. Fish Biol J Medaka 5: 11–15

Packer A (2001) Medaka on the move. Nat Genet 28: 302

Pohlman RF, Fedoroff NV, Messing J (1984) The nucleotide sequence of the maize controlling element Activator. Cell 37: 635–643

Ponnazhagan S, Hou L, Kwon BS (1994) Structural organization of the human tyrosinase gene and sequence analysis and characterization of its promoter region. J Invest Dermatol 102: 744–748

Rubin E, Lithwick G, Levy AA (2001) Structure and evolution of the *hAT* transposon superfamily. Genetics 158: 949–957

Sakaizumi M, Moriwaki K, Egami N. (1983) Allozymic variation and regional differentiation in wild populations of the fish *Oryzias latipes*. Copeia 2: 311–318

Tomita H (1975) Mutant genes in the medaka. In: Medaka (Killifish): Biology and Strains, T. Yamamoto, ed. Yugakusha Publ.,

Tokyo, pp 251–272
Tomita H (1992) The lists of the mutants and strains of the medaka, common gambusia, silver crucian carp, goldfish, and golden venus fish maintained in the Laboratory of Freshwater Fish Stocks, Nagoya University. Fish Biol J Medaka 4: 45–47
Wakamatsu Y, Pristyazhnyuk S, Kinoshita M, Tanaka M, Ozato K (2001) The see-through medaka: a fish model that is transparent throughout life. Proc Natl Acad Sci 98: 10046–10050

(Received October 28, 2001 / Invited Review)