

Biochemical and Biophysical Research Communications 293 (2002) 327-331



# Conserved physical linkage of *GnRH-R* and *RBM8* in the medaka and human genomes

Kataaki Okubo,<sup>a</sup> Hiroshi Mitani,<sup>b</sup> Kiyoshi Naruse,<sup>c</sup> Mariko Kondo,<sup>b</sup> Akihiro Shima,<sup>b</sup> Minoru Tanaka,<sup>d</sup> and Katsumi Aida<sup>a,\*</sup>

Department of Aquatic Bioscience, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Bunkyo, Tokyo 113-8657, Japan
Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba 277-8652, Japan
Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Bunkyo, Tokyo 113-0033, Japan
Division of Biological Sciences, Graduate School of Science, Hokkaido University, Kita, Sapporo 060-0810, Japan

Received 17 February 2002

#### Abstract

Candidate genes for human type II gonadotropin-releasing hormone receptor (*GnRH-RII*) reside on two separate loci, 1q12-q21 and 14q21-23, yet neither locus generates functional GnRH-RII. Instead, their opposite DNA strands encode functional RNA-binding motif protein 8 (RBM8s), which is also encoded by another locus, 5q13-q14. To elucidate the mechanism through which such multiple human *GnRH-RII/RBM8* loci arose, here we have defined an *RBM8* locus in a comparative model species, the medaka *Oryzias latipes*. The medaka *RBM8*, which exists as a single copy gene, is linked to, but does not overlap with, *GnRH-R2* on linkage group (LG) 16, demonstrating the ancient origin of the physical linkage between *GnRH-R* and *RBM8*. The medaka LG 16 contains orthologous segments to the human chromosome 1 and therefore the 1q12-q21 locus would be an originating human *GnRH-RII/RBM8* segment. Furthermore, like the human *RBM8s* on 1q12-q21 and 5q13-q14 but not that on 14q21-q23, the medaka *RBM8* is a multiexon gene, indicating that the 14q21-q23 and 5q13-q14 loci were generated by retrotransposition and segmental genomic duplication, respectively, of the originating 1q12-q21 locus. © 2002 Elsevier Science (USA). All rights reserved.

Keywords: Gonadotropin-releasing hormone receptor (GnRH-R); RNA-binding motif protein 8 (RBM8); Linkage group; Chromosome; Mapping; Medaka

Gonadotropin-releasing hormone (GnRH) plays pivotal roles in the neural control of vertebrate reproduction [1–4]. There is increasing evidence that a single vertebrate species possesses multiple GnRH receptor (GnRH-R) subtypes [5–9]. In humans, however, only one functional GnRH-R, designated type I GnRH-R, has been identified to date [10,11]. The human genomic sequences for an alternative GnRH-R, designated type II GnRH-R (GnRH-RII), were identified recently on two separate loci, 1q12-q21 and 14q21-q23, yet both two copies appear vestigial [12–14]. Instead, their opposite DNA strands encode the RNA-binding motif protein 8 (RBM8), a recently identified novel member of a highly conserved RBM family [12–14]. RBM8 is ubiquitously expressed in various tissues and has been implicated in

To help answer this question, here we have taken a comparative approach, utilizing the medaka *Oryzias latipes*, a well-defined genomic model species for which increasing amount of genetic linkage information is available [17]. We have isolated and characterized the medaka *RBM8* and examined the evolution of *GnRH-R/RBM8* loci in humans.

## Materials and methods

Isolation of the cDNA for the medaka RBM8. Double-strand cDNA was synthesized from whole brain of a medaka strain, the

germ plasm development in embryo, breast, and ovarian tumor suppression [15,16]. It is noteworthy that in addition to the 1q12-q21 and 14q21-q23, *RBM8* resides on 5q13-q14 [15,16]. The question therefore arises on the mechanism through which such multiple *GnRH-RII/RBM8* loci arose in humans.

<sup>\*</sup> Corresponding author. Fax: +81-3-5841-5288. E-mail address: aida@uf.a.u-tokyo.ac.jp (K. Aida).

Table 1 Primers used in this study

Name	Sequence (5′–3′)	nt Position	
F1	GTIGARCARGAYGGIGAYGARCC	171–193	
F2	GCTACGCACTGGTGGAGTATGAGACG	346–371	
F3	ATGGAGGGCTCAACGCTCAGGATC	393-417	
F4	GCCTCAGAGATCTGTAGAAGGCTG	197–220	
F5	TGGCGGACGTGTTGGACCTTCATGA	10–34	
F6	CGGTCCAGAGCCAAAGAGGACTATG	141–165	
F7	GGAAGACATCCACGATAAGTTCTCAGAG	260–287	
F8	CGCACTGGTGGAGTATGAGACGTAC	350–374	
F9	TCACGCTCAGCCCAGACTGCGT	1st Intron	
R1	TCCACRCTDATIGGYTGICCCAT	420-442	
R2	CCAAACTCTGAGAACTTATCGTGGATGTC	264–292	
R3	AGGATCCAGCCTTCTACAGATCTCTGAG	199–226	
R4	TCCAACTATCACGCGCCAAAGATGGTG	532–558	
R5	CAGGCTCATCTCCGTCCTGTTCCA	172–195	
R6	CCCTTCAGATAGCCAGTTCTGCGATC	321–346	
R7	ATCCTGAGCGTTGAGCCCCTCCAT	393–416	
R8	AAATGGAATAAAATCCAGGTCACAAAGAGC	760–789	
R9	GACCGCCTTTCGTTGAGTTAGCTG	1st Intron	

himedaka, as previously described [18]. A degenerate primer pair, F1/R1 (Table 1), was designed based on conserved segments among the human RBM8s (accession nos. AF198620, AF231511, and AF231512) and their candidate orthologs were found in the Xenopus laevis (AW200013) and zebrafish Danio rerio (AI943400) EST databases. The brain cDNA was amplified with the F1/R1 primer pair. The PCR product was analyzed on a 2.0% agarose gel and ligated into pBluescript SK(-) (Stratagene, La Jolla, CA). The purified plasmids were sequenced using a DNA sequencer ABI PRISM 310 (Perkin-Elmer, Branchburg, NJ). Sequence data were analyzed with Sequencer software version 3.1.1 (Hitachi, Tokyo, Japan) and SeqEd software version 1.0.3 (Perkin-Elmer). The nucleotide sequence was determined by analyzing more than three clones from distinct amplifications to avoid PCR errors. Rapid amplification of cDNA ends (RACE) was subsequently carried out. Gene-specific primers, F2/F3 and R2/R3 (Table 1), were designed to be used as primary/nested primers for 3'- and 5'-RACE, respectively. Electrophoresis, subcloning, sequencing, and sequence analysis were performed as described above. Amino acid identity between the medaka and human RBM8 proteins was calculated using Mac Vector version 6.0 (Oxford Molecular, Beaverton, OR).

Northern blot analysis. Ten  $\mu$ g of poly(A)<sup>+</sup> RNA from the brain, liver, and gonad (the ovary and testis were mixed together) of the himedaka was subjected to electrophoresis on a 0.9% agarose gel and transferred to a nylon membrane (Hybond-N<sup>+</sup>; Amersham Pharmacia Biotech, Buckinghamshire, UK). Northern blot analysis was performed using Alkphos Direct System (Amersham Pharmacia Biotech) according to manufacturer's instructions. A 362 bp cDNA probe covering 3'-half of the *RBM8* open reading frame was generated by PCR amplification with a primer pair, F4/R4 (Table 1). The membrane was exposed to Hyperfilm ECL (Amersham Pharmacia Biotech) at room temperature for 30 min.

Isolation of the gene for the medaka RBM8. The transcribed region of the medaka RBM8 was determined by means of PCR amplifications of the corresponding genomic regions and alignment of the contigs by sequence analysis. The following primer pairs were used for the PCR amplifications: F5/R5, F6/R6, F7/R7, and F8/R8 (Table 1). PCR was performed using LA Taq DNA polymerase (Takara). Electrophoresis, subcloning, sequencing, and sequence analysis were performed as described above.

Genomic Southern blot analysis. Ten  $\mu g$  of the himedaka genomic DNA was digested separately with restriction enzymes AccI, HincII, NheI, SpeI, SphI, and StuI and electrophoresed on a 0.8% agarose gel.

Genomic Southern blot analysis was carried out in the same way as described in Northern blot analysis.

Chromosomal mapping. An insertion/deletion polymorphism between two inbred medaka strains, HNI and AA2, was found within the first intron of *RBM8*. A primer pair F9/R9 was designed to amplify the polymorphism-containing sequences (Table 1). PCR amplifications of the HNI and AA2 genomic DNAs yielded DNA fragments of 93 and 78 bp, respectively. Chromosomal mapping of *RBM8* was performed using reference-typing DNA panels derived from 39 offsprings of a backcross between a HNI/AA2 male F1 and an AA2 female parental line [17]. Genotypes were analyzed by amplification of the polymorphic DNA region followed by 10% polyacrylamide gel electrophoresis.

#### Results

Isolation and characterization of the cDNA for the medaka RBM8

The medaka RBM8 cDNA isolated in the current study (accession no. AB069905) consists of 832 bp and encodes RBM8 of 173 amino acids (Fig. 1). The medaka RBM8 protein contains a highly conserved sequence of RNA-recognition motif (Fig. 1), which is also found in the human RBM8s [13–16]. The predicted amino acid sequence of the medaka RBM8 exhibited significant identity (86–88%) to the human RBM8s. The expression pattern of the medaka RBM8 was subsequently evaluated by Northern blot analysis using the brain, liver, and gonad. A single transcript species of approximately 1.0 kb was detected in each of the three tissues (Fig. 2).

Isolation and characterization of the gene for the medaka RBM8

Sequencing of the medaka *RBM8* coding region (accession no. AB069906) revealed that it spread over six

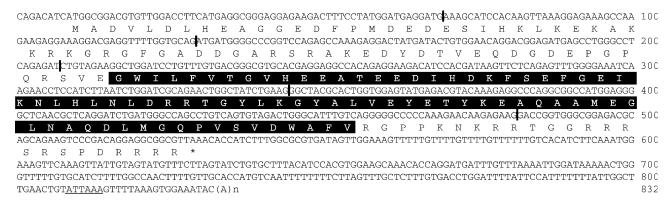


Fig. 1. Nucleotide and deduced amino acid sequences of the medaka RBM8 cDNA. Boxed amino acids indicate an RNA-recognition motif. The stop codon is denoted by an asterisk and the polyadenylation signal is underlined. Five intron positions are indicated by thick, short lines.

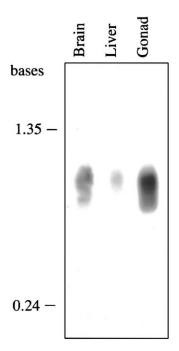


Fig. 2. Northern blot analysis for the medaka RBM8 using the brain, liver, and gonadal poly(A)<sup>+</sup> RNA. A single transcript species of approximately 1.0 kb was detected in each of the three tissues.

exons, covering more than 5 kb of genomic DNA (Fig. 3). The sequences at the exon/intron junctions all follow the GT/AG rule for consensus splice donor and acceptor sites. Genomic Southern blot analysis was carried out to determine the copy number of the medaka *RBM8*. Hybridization with the RBM8 cDNA probe yielded 1, 2, 1, 1, 1, and 2 bands for *AccI*, *HincII*, *NheI*, *SpeI*, *SphI*, and *StuI* digestions, respectively (Fig. 4). This hybridization pattern is consistent with the *RBM8* restriction map (data not shown), indicating that the medaka *RBM8* is a single copy gene. The medaka *RBM8* was then mapped using interspecific backcross panels. The medaka has 24 chromosomes per haploid set and several hundreds of genetic markers in this teleost species have been classified into 24 linkage groups (LGs), demonstrating that



Fig. 3. Schematic diagram illustrating the organization of the medaka *RBM8*. Boxes represent exons and horizontal lines adjacent to exons represent introns. Filled and open boxes show open reading frame and untranslated region, respectively. Size in bp is indicated.

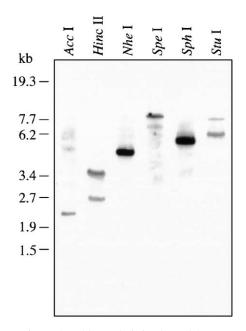


Fig. 4. Genomic Southern blot analysis for the medaka *RBM8*. A single signal was obtained from the genomic DNA digested with *Acc*I, *Nhe*I, *Spe*I, and *Sph*I. Two signals were obtained from the *Hin*cII and *Stu*I digestions. The cDNA probe used contained one restriction site for each of *Hin*cII and *Stu*I.

one LG represents one chromosome [17]. *RBM8* was mapped on LG 16, where *GnRH-R2*has previously been assigned [9] (Fig. 5). Out of 39 panels, five showed recombination between *RBM8* and *GnRH-R2*. Independent confirmation for the assignment of *RBM8* on LG

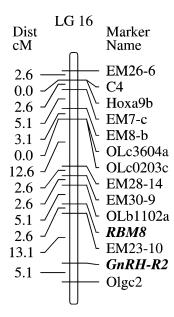


Fig. 5. The location of *RBM8* in relation to its linked genes including *GnRH-R2* on LG 16. Recombination distances between loci in centimorgans (cM) are shown on the left.

16 was provided by analysis of other backcross panels derived from 52 offsprings between a HNI/AA2 female F1 and an AA2 male parental line (data not shown). In addition, using the 39 backcross panels described above, we identified the following three genes which are located on LG 16 and are orthologous to human genes on chromosome 1: *OLc3604a*, *OLb1102a*, and *OLc0203c* (Table 2).

### Discussion

Our interest in the presence of multiple *RBM8/GnRH-RII* loci in the human genome led us to characterize the corresponding region of a comparative model species, the medaka. The medaka *RBM8* encodes a protein with >86% amino acid identity to its human counterparts and the medaka transcript was detected in all three tissues analyzed here, which coincides with the previous reports showing the ubiquitous expression of the human *RBM8s* [13–16]. Thus the medaka *RBM8* identified in the current

study should be a functional medaka ortholog of the human *RBM8s*.

The physical linkage of GnRH-R and RBM8 found on the human chromosomes 1q12-q21 and 14q21-q23 has also been identified on the medaka LG 16. The physical linkage of RBM8/GnRH-R would therefore date from the last common ancestor of teleosts and mammals and has been conserved throughout vertebrate evolution. The medaka LG 16 also contains *OLc3604a*, OLb1102a, and OLc0203c, whose orthologous genes have been assigned on the human chromosome 1, showing that the medaka LG 16 shares conserved synteny with the human chromosome 1. The medaka LG 16 would therefore contain an orthologous segment of the human chromosome 1. These lines of evidence indicate that the 1q12-q21 locus is an originating GnRH-RII/RBM8 segment in humans had and GnRH-RII/ *RBM8* on 14q21-q23 and *RBM8* on 5q13-q14 are its duplicates. The human RBM8s on 1q12-q21 and 5q13q14 are intron-containing genes, whereas that on 14q21q23 lacks introns [14,16]. Here we have demonstrated that the exon/intron structure of the medaka RBM8 is identical to the human counterparts on 1q12-q21 and 5q13-q14; each consists of six exons and five introns. Hence the RBM8 of the last common ancestor of the medaka and humans had a multiexon structure. Accordingly, the most likely origin of the intronless RBM8A on the human chromosome 14q21-q23 was via retrotransposition of a fully processed RBM8 mRNA derived from the originating 1q12-q21 locus. Thus, the 14g21-g23 locus contains only the RBM8A mRNA sequence but not the promoter region or full-length open reading frame for GnRH-RII and it is therefore reasonable that no transcript is produced from this locus [14]. On the other hand, the intron-containing *RBM8* on 5q13-q14 would arise from the 1q12-q21 locus by a chromosomal segment duplication. Furthermore, the fact that the three human RBM8s loci exhibit remarkably high sequence identities to one another indicates that these duplicates arose within the mammalian, perhaps primate, lineage.

The present observation on the medaka, however, differs from the case in humans in that the medaka *RBM8* and *GnRH-R2* do not overlap each other, whereas the antisense DNA strands of the human

Table 2 List of medaka genes on LG 16 whose human orthologs reside on chromosome 1  $\,$ 

Gene name	Human ortholog (location)	Primers used to amplify polymorphism-containing sequence	Detection of polymorphism	Accession no.
OLc3604a	LOC51118 (1p34.3)	5'-AATCATCAGCCCTAACAAAACCATCTCG-3' 5'-CCTCCTCCTGTCCTTTCTTCGTGATATG-3'	MseI	AU241204
OLb1102a	HSPC003 (1)	5'-GCCCGTGTTTATTCAGTCAGATCA-3' 5'-TGTCGTAGGCGGCCCAGATGTT-3'	Insertion/deletion <i>Msp</i> I	AV669745
OLc0203c	CTPS (1p34.1)	5'-CAGAGGGACAGAAGGAAAGATTATGGCC-3' 5'-CAGTACGCTGCTGCATTTGAAAA-3'	Insertion/deletion	AU177595

GnRH-RIIs encode 3'-untranslated region of RBM8s [12–14]. The syntenic pattern of RBM8 and GnRH-R is also different between the medaka and humans in that the medaka RBM8 is linked to GnRH-R2, which is predicted to be paralogous, but not orthologous, to the human *GnRH-RII*. Its medaka ortholog is considered to be GnRH-R1, which resides on LG 3 [9]. Genomic Southern blot analysis revealed the presence of a single copy RBM8 in the medaka genome, indicating that no additional RBM8 remains to be found in the medaka. Thus it can be concluded that the medaka RBM8 is linked to GnRH-R2, but not GnRH-R1. These data suggest that the duplication of an ancestral GnRH-R that gave rise to multiple GnRH-R paralogs in vertebrates occurred after the formation of a physical linkage between RBM8 and GnRH-R. It therefore seems possible that GnRH-R1-linked RBM8 was also present in the medaka progenitor genome, but it may have been lost during evolution.

#### References

- H. Matsuo, Y. Baba, R.M. Nair, A. Arimura, A.V. Schally, Structure of the porcine LH- and FSH-releasing hormone. I. The proposed amino acid sequence, Biochem. Biophys. Res. Commun. 43 (1971) 1334–1339.
- [2] M. Amoss, R. Burgus, R. Blackwell, W. Vale, R. Fellows, R. Guillemin, Purification, amino acid composition and N-terminus of the hypothalamic luteinizing hormone releasing factor (LRF) of ovine origin, Biochem. Biophys. Res. Commun. 44 (1971) 205–210
- [3] R.L. Moss, S.M. McCann, Induction of mating behavior in rats by luteinizing hormone releasing factor, Science 181 (1973) 177– 179.
- [4] D.W. Pfaff, Luteinizing hormone-releasing factor potentiates lordosis behavior in hypophysectomized ovariectomized female rats, Science 182 (1973) 1148–1149.
- [5] B. Troskie, N. Illing, E. Rumbak, Y.M. Sun, J. Hapgood, S. Sealfon, D. Conklin, R. Millar, Identification of three putative GnRH receptor subtypes in vertebrates, Gen. Comp. Endocrinol. 112 (1998) 296–302.
- [6] L. Wang, J. Bogerd, H.S. Choi, J.Y. Seong, J.M. Soh, S.Y. Chun, M. Blomenrohr, B.E. Troskie, R.P. Millar, W.H. Yu, S.M. McCann, H.B. Kwon, Three distinct types of GnRH receptor characterized in the bullfrog, Proc. Natl. Acad. Sci. USA 98 (2001) 361–366.

- [7] J.D. Neill, L.W. Duck, J.C. Sellers, L.C. Musgrove, A gonadotropin-releasing hormone (GnRH) receptor specific for GnRH II in primates, Biochem. Biophys. Res. Commun. 282 (2001) 1012– 1018.
- [8] R. Millar, S. Lowe, D. Conklin, A. Pawson, S. Maudsley, B. Troskie, T. Ott, M. Millar, G. Lincoln, R. Sellar, B. Faurholm, G. Scobie, R. Kuestner, E. Terasawa, A. Katz, A novel mammalian receptor for the evolutionarily conserved type II GnRH, Proc. Natl. Acad. Sci. USA 98 (2001) 9636–9641.
- [9] K. Okubo, S. Nagata, R. Ko, H. Kataoka, Y. Yoshiura, H. Mitani, M. Kondo, K. Naruse, A. Shima, K. Aida, Identification and characterization of two distinct gonadotropin-releasing hormone receptor subtypes in a teleost, the medaka *Oryzias latipes*, Endocrinology 142 (2001) 4729–4739.
- [10] S.S. Kakar, L.C. Musgrove, D.C. Devor, J.C. Sellers, J.D. Neill, Cloning, sequencing, and expression of human gonadotropin releasing hormone (GnRH) receptor, Biochem. Biophys. Res. Commun. 189 (1992) 289–295.
- [11] L. Chi, W. Zhou, A. Prikhozhan, C. Flanagan, J.S. Davidson, M. Golembo, N. Illing, R.P. Millar, S.C. Sealfon, Cloning and characterization of the human GnRH receptor, Mol. Cell. Endocrinol. 9 (1993) R1–R6.
- [12] R. Millar, D. Conklin, C. Lofton-Day, E. Hutchinson, B. Troskie, N. Illing, S.C. Sealfon, J. Hapgood, A novel human GnRH receptor homolog gene: abundant and wide tissue distribution of the antisense transcript, J. Endocrinol. 162 (1999) 117–126.
- [13] D.C. Conklin, M.W. Rixon, R.E. Kuestner, M.F. Maurer, T.E. Whitmore, R.P. Millar, Cloning and gene expression of a novel human ribonucleoprotein, Biochem. Biophys. Acta 1492 (2000) 465–469.
- [14] B. Faurholm, R.P. Millar, A.A. Katz, The genes encoding the type II gonadotropin-releasing hormone receptor and the ribonucleoprotein RBM8A in humans overlap in two genomic loci, Genomics 78 (2001) 15–18.
- [15] X.F. Zhao, N.J. Nowak, T.B. Shows, P.D. Aplan, MAGOH interacts with a novel RNA-binding protein, Genomics 63 (2000) 145, 148
- [16] A.M. Salicioni, M. Xi, L.A. Vanderveer, B. Balsara, J.R. Testa, R.L. Dunbrack Jr., A.K. Godwin, Identification and structural analysis of human RBM8A and RBM8B: two highly conserved RNA-binding motif proteins that interact with OVCA1, a candidate tumor suppressor, Genomics 69 (2000) 54–62.
- [17] K. Naruse, S. Fukamachi, H. Mitani, M. Kondo, T. Matsuoka, S. Kondo, N. Hanamura, Y. Morita, K. Hasegawa, R. Nishigaki, A. Shimada, H. Wada, T. Kusakabe, N. Suzuki, M. Kinoshita, A. Kanamori, T. Terado, H. Kimura, M. Nonaka, A. Shima, A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution, Genetics 154 (2000) 1773–1784.
- [18] K. Okubo, M. Amano, Y. Yoshiura, H. Suetake, K. Aida, A novel form of gonadotropin-releasing hormone in the medaka, *Oryzias latipes*, Biochem. Biophys. Res. Commun. 276 (2000) 298–303.