Gene Expression Patterns 12 (2012) 46–52

Contents lists available at SciVerse ScienceDirect

Gene Expression Patterns

journal homepage: www.elsevier.com/locate/gep

Molecular cloning and expression of the *col2a1a* and *col2a1b* genes in the medaka, *Oryzias latipes*

Tomohiro Matsumoto [a,b,c,1], Tomonori Deguchi [a,*,1], Takashi Kawasaki [a], Shunsuke Yuba [a], Junichi Sato [b,c]

[a] Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 3-11-46 Nakouji, Amagasaki, Hyogo 661-0974, Japan
[b] First Department of Oral and Maxillofacial Surgery, Tsurumi University, School of Dental Medicine, 2-1-3 Tsurumi, Tsurumi-ku, Yokohama, Kanagawa 230-0063, Japan
[c] Division of Oral and Maxillofacial Implantology, Tsurumi University, 2-1-3 Tsurumi, Tsurumi-ku, Yokohama, Kanagawa 230-8501, Japan

ARTICLE INFO

Article history:
Received 28 February 2011
Received in revised form 17 October 2011
Accepted 4 November 2011
Available online 22 November 2011

Keywords:
col2a1a
col2a1b
Medaka
Cartilage
Notochord
Floor plate
Hypochord
Otic vesicle
Pectoral fin
Craniofacial cartilage

ABSTRACT

The *Col2a1* gene is expressed in notochord, otic vesicle, cartilaginous tissue and the anlage of endochondral bone during development in higher vertebrates. Type II collagen, a homotrimeric product of the *Col2a1* gene, functions as a key regulatory protein for cartilage development and endochondral ossification. In medaka and zebrafish, a single homolog of the *col2a1* gene has been identified. However, it is necessary to note that many genes are duplicated in teleost fishes. To clarify function of *col2a1* genes in teleost fishes and to further understand the process of cartilage development and endochondral ossification, we cloned and mapped the gene loci of two *col2a1* orthologs in medaka. The proteins encoded by both medaka *col2a1a* and *col2a1b* genes were highly conserved (85.3% and 82.6%) relative to human COL2A1, but synteny was not observed. We also examined the expression patterns of *col2a1a* and *col2a1b* during embryonic development. Whole-mount *in situ* hybridization data suggests that expression patterns of both medaka *co2a1a* and *col2a1b* genes are similar to that of zebrafish *co2a1* in the early embryonic stages. In medaka, the two *col2a1* genes show a closely correlated pattern of spatial and temporal expression. In late embryonic stages, however, there were differences in both expression patterns in the pectoral fin. This study is the first report of two homologs of *col2a1* in teleosts and also the first examination of *col2a1a* and *col2a1b* expression patterns in this group.

In mammals, there are two main mechanisms of bone formation, endochondral ossification, and intramembranous ossification. Endochondral ossification is the major process of skeletal development through the cartilaginous anlage, which is subsequently mineralized and then replaced by bone. Type II collagen, a homotrimeric product of the *Col2a1* gene, was reported to function as a key regulatory protein for endochondral ossification and cartilage development (Hall, 1975; Cheah et al., 1991; Ng et al., 1993; Helminen et al., 1993; Reddi, 1994; Li et al., 1995). During chondrogenesis, *Sox9*, which is known as transcription factor regulating the morphogenesis and differentiation of cartilage and bone, is coexpressed with *Col2a1*. *Sox9* has been shown to directly activate transcription of *Col2a1* by binding to an enhancer in its first intron (Bell et al., 1997; Lefebvre et al., 1997; Ng et al., 1997; Bi et al., 1999; Akiyama et al., 2002). Furthermore, transient expression of *Col2a1* has been observed in non-chondrogenic tissues such as the notochord, heart, brain, eye, and otic capsule during mouse embryogenesis (Cheah et al., 1991; Ng et al., 1993; Maddox et al., 1998). While small teleost fishes such as medaka and zebrafish are useful model animals in clarifying the mechanism of skeletal development (Yan et al., 1995, 2002; Sachdev et al., 2001; Renn et al., 2006; Inohaya et al., 2007), the precise mechanism of endochondral ossification has not been investigated (Inohaya et al., 2007). In zebrafish, the *col2a1* gene has been previously identified and its role in development has been studied (Yan et al., 1995). Although *col2a1* has been used as a marker of chondrogenic tissues in early embryonic development (Yan et al., 2002; Lang et al., 2006), *col2a1* function in late embryonic development has not been well studied. Furthermore, there are often two gene orthologs in teleost fishes due to gene duplication (Kasahara et al., 2007). Such gene duplications can enable novel functions to evolve. However, duplication of *col2a1* genes in teleost fishes has not been investigated. We therefore searched for orthologs of the *col2a1* gene in medaka and consequently cloned them. Expression patterns of these mRNAs were examined by whole-mount *in situ* hybridization during development up to the fry stage.

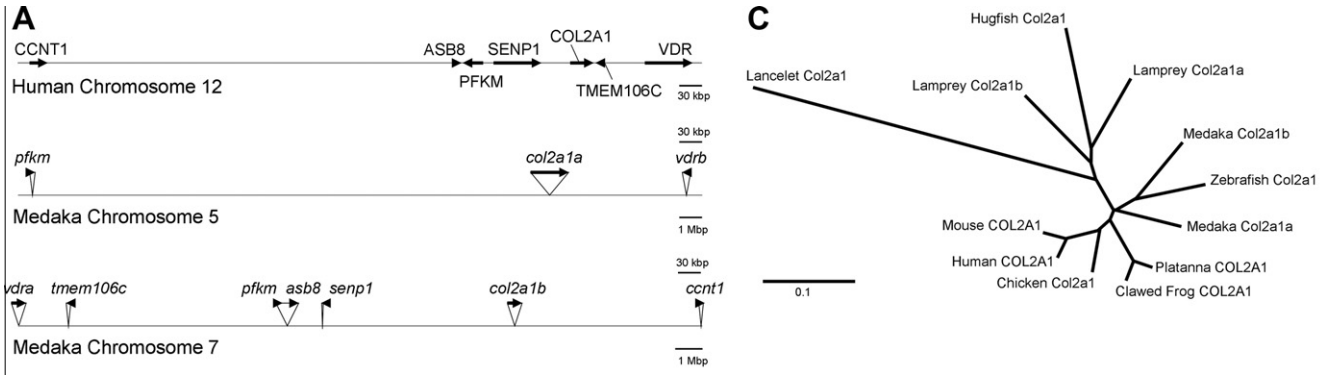
1. Results

1.1. Medaka col2a1a and col2a1b genes

We predicted the position and sequence of the *col2a1* orthologs in medaka based on an *in silico* analysis of the medaka genome

* Corresponding author. Tel.: +81 727 51 9834; fax: +81 66 494 7861.
E-mail address: tomonori-deguchi@aist.go.jp (T. Deguchi).
[1] These authors contributed equally to this work.

1567-133X/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.gep.2011.11.002



B

```
Human COL2A1      1:MIRLGAPQTLVLLT---LLVAAVLRCQGQD-VQEAGSCVQDGQRYNDKDVWKPEPCRICVCDTGTVLCDDIICEDVKDCLSPEIPFGECCPICPTDLATASGQPGPKGQKGEPGDIKDIVGPKGPPGPQGPAGEQGPRGDRGDKGEKGAP 146
Mouse COL2A1      1:.......S.....---..I.....-A.....L.N....K.....SS.....N.....P-...N.....A.....KL.....R..I..R..... 145
Zebrafish Col2a1  1:..F..LDSR..L..VATHSVLLSLV...QE.DQE.F.G.....Q.A.RA.....V....S.....EV....LN..AN.I.SP.....A.TDDPI.SL.A.....T.V...R..A..M..P....T..E..A.....S. 150
Medaka Col2a1a    1:..SFLESR.VL..VATQ.CLL..V...LD.DLEDEF.....I.....F.SLV....T.VEL...PK.....S.EP.SP.V..I.....T.V...R..S..M..P.....P..E.....SL 150
Medaka Col2a1b    1:..FS.VDSR.VL..VASQVVLLS.V...ED.--ED..A.I..E.Q.YN.....S.A....E....EI.E.PN.V..S.....A.ASAPIERI.A.....VANV.....M..P.....EA.A..D..N. 148

Human COL2A1      147:GPRGRDGEPGTPGNPGPPGPPGPPGPPGLG-GNFAAQMAGGFDEKAGGAQLGVMQGPMGPMGPRGPPGPAGAPGPQGFQGNPGEPGEPGVSGPMGPRGPPGPPGKPGDDGEAGKPGKAGERGPPGPQGARGFPGTPGLPGVKGHRGYPGL 295
Mouse COL2A1      146:.....A.....SA.....Y.....M.....A.....S....L..... 295
Zebrafish Col2a1  151:.....N.....-.....M.....S.....T...PA.AL.....S.....I....H.. 299
Medaka Col2a1a    151:.....N.....-.....A..M.....T.S.....Q.....S....S.....I..... 299
Medaka Col2a1b    149:.....-.....S.....M.....S.....A.....S.PA.....S....S...S.....P.....A.....I....H.. 297

Human COL2A1      296:DGAKGEAGAPGVKGESGSPGENGSPGPMGPRGLPGERGRTGPAGAAGARGNDGQPGPAGPPGPVGPAGGPGFPGAPGAKGEAGPTGARGPEGAQGPRGEPGTPGSPGPAGASGNPGTDGIPGAKGSAGAPGIAGAPGFPGPRGPPGPQGA 445
Mouse COL2A1      296:.....S.....N..... 445
Zebrafish Col2a1  300:.....A.A...A..N..S.A.....P.AT.....L.....A.....S..S.....A.....S..... 449
Medaka Col2a1a    300:.....T..V.A...A.A...S.A.....P.ST.....L.....A.....S..S.....S.....S..... 449
Medaka Col2a1b    298:.....T..A.A...A.AS....A.....P.A.....L.....A.....S.....L..A..P....R.A..... 447

Human COL2A1      446:TGPLGPKGQTGEPGIAGFKGEQGPKGEPGPAGPQGAPGPAGEEGKRGARGEPGGVGPIGPPGERGAPGNRGFPGQDGLAGPKGAPGERGPSGLAGPKGANGDPGRPGEPGLPGARGLTGRPGDAGPQGKVGPSGAPGEDGRPGPPGPQGA 595
Mouse COL2A1      446:.....A.....D.....T.....A.....T..... 595
Zebrafish Col2a1  450:.....S.D...P.....A.....R.VL....P...S.....P.....SA..L.....A.....D..VP..S....GT.....A.....AT.....L.. 599
Medaka Col2a1a    450:.....TS.D.....A.....I..P.L.....Q.....P.....AA..L.....Q.....P.AS.A...T.....S..... 599
Medaka Col2a1b    448:.....S.D..LP.....V.....L..V.....P.....TA..L.....P.....A.....D..VA.AS.....T.....T.....T..S..... 597

Human COL2A1      596:RGQPGVMGFPGPKGANGEPGKAGEKGLPGAPGLRGLPGKDGETGAAGPPGPAGPAGERGEQGAPGPSGFQGLPGPPGPPGEGGKPGDQGVPGEAGAPGLVGPRGERGFPGERGSPGAQGLQGPRGLPGTPGTDGPKGASGPAGPPGAQGP 745
Mouse COL2A1      596:.....A.....S.....Q....I.....A..D..... 745
Zebrafish Col2a1  600:.....P.....V.RT.....PS.....V.AV.....Q.....T.A...P.....G..A.PT.....GA.P.....I....AA..... 749
Medaka Col2a1a    600:.....T.....S.....L.....S.....Q.....A.AT.....PA.P.....I....SR..... 749
Medaka Col2a1b    598:.....P.....V.R.....S.....Q.....A.....A.....S.....A..Q.....GA.P.....S.....I....AA.P... 747

Human COL2A1      746:PGLQGMPGERGAAGIAGPKGDRGDVGEKGPEGAPGKDGGRGLTGPIGPPGPAGANGEKGEVGPPGPAGSAGARGAPGERGETGPPGPAGFAGPPGADGQPGAKGEQGEAGQKGDAGAPGPQGPSGAPGPQGPTGVTGPKGARGAQGPPGA 895
Mouse COL2A1      746:.....S..T.....D..... 895
Zebrafish Col2a1  750:.....V..S.A.....S.....S.....S.P..A...T..I.SI.AP.....D...I.A.....N.....S.....S.....V.....A... 899
Medaka Col2a1a    750:Q.....P.....L.....A....A.....G.P.....S..A..S.AP.T.....D.....T.....P....E.....A.....S..... 899
Medaka Col2a1b    748:.....G..P.A.....N.Q.....A.....P..A...T.ST..I.AP.V.....D...V.....L..S.....ES.....V.....S.....A... 897

Human COL2A1      896:TGFPGAAGRVGPPGSNGNPGPPGPPGPSGKDGPKGARGDSGPPGRAGEPGLQGPAGPPGEKGEPGDDGPSGAEGPPGPQGLAGQRGIVGLPGQRGERGFPGLPGPSGEPGKQGAPGASGDRGPPGPVGPPGLTGPAGEPGREGSPGADGP 1045
Mouse COL2A1      896:.....A.....A.....A.....V.....D.....A.....LD..... 1045
Zebrafish Col2a1  900:.....P.....AA..A.....V...A.....DA..R..P.A.....A.E...P.PD..S..A.....G.....T....N..S... 1049
Medaka Col2a1a    900:.....P.....A..A..P.....V...P..A..Q.DA..R....A.....DA.E...N.PS..S.....GP.....N..S... 1049
Medaka Col2a1b    898:.....P.....A..A.AA.....T...A.....P.DA..R.....EE.KP.TD.....S.....G.....T..S... 1047

Human COL2A1      1046:PGRDGAAGVKGDRGETGAVGAPGAPGPPGSPGPAGPTGKQGDRGEAGAQGPMGPSGPAGARGIQGPQGPRGDKGEAGEPGERGLKGHRGFTGLQGLPGPPGPSGDQGASGPAGPSGPRGPPGPVGPSGKDGANGIPGPIGPPGPRGRSGE 1195
Mouse COL2A1      1046:.....L.....A.....S..Q.....S..... 1195
Zebrafish Col2a1  1050:.....E..N..PI.....A..A..SV..I.....N.P...A..P.....MV.....A....Q.....SP.....A.....AK..S.....A....S..Q..... 1199
Medaka Col2a1a    1050:.....S..I.....N..P...S.....A..A...V.....P.P...A.....MP.....G....Q.....QA.....T.....A.Q.....R.....L.....G.D 1199
Medaka Col2a1b    1048:.....T.I..E..N..PA.....A..A...V..L.....N.....A..P.....MA.....S....Q.....A..S.PA.....AK.....A.....N..QA..... 1197

Human COL2A1      1196:TGPAGPPGNPGPPGPPGPPGPGIDMSAFAGLGPREKGPDPLQYMRADQAAG-GLRQHDAEVDATLKSLNNQIESIRSPEGSRKNPARTCRDLKLCHPEWKSGDYWIDPNQGCTLDAMKVFCNMETGETCVYPNPANVPKKNWWSSKSKEK 1344
Mouse COL2A1      1196:...V....S.....Q.....M.....E.DS-T.....V.....D.....Q.....T..R..... 1344
Zebrafish Col2a1  1200:S..V.....SQP.....R.....E.SS-S.....V.....I.G...D...D.....S.....V...L..SAA..I.....K.STPKI.R...T....AQ 1348
Medaka Col2a1a    1200:S.....QP..S...R.....AGN.....I.....N.....K.....D.....V..I.....S.....SSI.R.....R-N 1348
Medaka Col2a1b    1198:.....L.....QT..S...R.....E.SN-S.....V...S.....S...NM...D.TQ.....F.....L...A..I.....RISK.....NL 1346

Human COL2A1      1345:KHIWFGETINGGFHFSYGDDNLAPNTANVQMTFLRLLSTEGSQNITYHCKNSIAYLDEAAGNLKKALLIQGSNDVEIRAEGNSRFTYTALKDGCTKHTGKWGKTVIEYRSQKTSRLPIIDIAPMDIGGPEQEFGVDIGPVCFL 1487
Mouse COL2A1      1345:.....M.....G.....M.....A..... 1487
Zebrafish Col2a1  1349:..V....SM.....A.GSQT.S.TTI.LN.....AT.T.....V..M.Q.T.....I.L.....GV.E...K....Q.A.....KT.....M.....AD.....A.... 1491
Medaka Col2a1a    1349:..V....M.....N....S.....AI..N.....A...L.....V..M.AST.....L.....SVME.S..E.....M.....V.....AD.....V..... 1491
Medaka Col2a1b    1347:.....S.....Q.GPI..S..I.LN.....A...L.....M.Q.T.....VL.....SVME...KR.S.R.SN..F..KTM.....V.....AD..... 1489
```

Fig. 1. The cloned medaka *col2a1* orthologs, *col2a1a* and *col2a1b*. (A) Gene location of human *COL2A1* and medaka *col2a1* orthologs. The schematic drawing of human chromosome 12 around *COL2A1* gene was modified from NCBI (http://www.ncbi.nlm.nih.gov/). (B) The amino acid sequences of *COL2A1* orthologs were compared by multiple alignments. Conserved residues with human COL2A1 are shown as dots. (C) Phylogenetic tree of amino acid sequences using the NJ method under 1000-times bootstrap conditions. Accession numbers: human (NP_001835.3), mouse (NP_112440.2), chicken (NP_989757.1), clawed frog (NP_989220.1), platanna (NP_001081258.1), zebrafish (NP_571367.1), hugfish (ABG36940.1), lamprey (Col2a1a, ABB53637.2; Col2a1b, ABB5368.2), and lancelet (ABG36939.1).

database (http://dolphin.lab.nig.ac.jp/medaka/). Two candidates of medaka *col2a1* were obtained, and named *col2a1a* and *col2a1b*. *Col2a1b* was identical with a medaka *col2a* that had been partially cloned by Nemoto et al. (2008). The medaka *col2a1a* gene was located on chromosome 7 while *col2a1b* was located in chromosome 5. Synteny (the physical co-localization of genetic loci on the same chromosome) of the *col2a1* gene was studied. Human *COL2A1* is located on chromosome 12. *CCNT1*, *ASB8*, *PFKM*, and *SENP1* genes are found upstream of *COL2A1*. *TMEM106C* and *VDR* are downstream of *Col2A1*. This synteny is well conserved between humans and mice. We also predicted the orthologs of these genes from an *in silico* analysis. The only candidates for medaka *pfkm* and *vdr*