Predicted Protein Structure of Medaka FoxA3 and Its Expression in Polster

Authors: Okamoto, Hiroyuki M., Nakayama, Ichiro, Nagoya, Hiroyuki, and Araki, Kazuo

Source: Zoological Science, 18(6) : 823-832

Published By: Zoological Society of Japan

URL: https://doi.org/10.2108/zsj.18.823

ZOOLOGICAL SCIENCE **18**: 823–832 (2001)

Predicted Protein Structure of Medaka *FoxA3* and Its Expression in Polster

Hiroyuki M. Okamoto*, Ichiro Nakayama, Hiroyuki Nagoya and Kazuo Araki

Cell Engineering Section, National Research Institute of Aquaculture, 224-1 Tamaki, Mie 519-0423, Japan

ABSTRACT—The forkhead box genes (Fox genes) are expressed in many tissues during embryogenesis and their products play a key role in organogenesis. We cloned two Fox genes from an embryonic cDNA library of medaka (*Oryzias latipes*). One was *MeHNF3β,* a homologue of *HNF3β* (FoxA2). The other was *Mefkh1* (medaka-forkhead1), related to *HNF3γ* (FoxA3). We found that the expression pattern of FoxA3 differs in some aspects between medaka and zebrafish. *Mefkh1* is expressed in the dorsal marginal zone, embryonic shield, polster, gut and slightly axial mesoderm, whereas zebrafish FoxA3 is expressed in the yolk syncytial layer (YSL), axial mesoderm, polster, gut and posterior neural crest cells (Odenthal and Nüsslein-Volhard, 1998, Dev Genes Evol 208, 245–258). In zebrafish, the level of expression of FoxA3 in the axial mesoderm is similar to that of FoxA2 (*axial*); in medaka, FoxA3 (*Mefkh1*) expression is mainly detected in the polster, and FoxA2 (*MeHNF3β*) is expressed in the axial mesoderm. This indicates that the combined functioning of FoxA2 and FoxA3 in the organogenesis of the polster and axial mesoderm differs between medaka and zebrafish.

INTRODUCTION

Inductive interaction is of fundamental importance to the development of multicellular organisms. We consider that the molecular mechanisms underlying inductive interactions are conserved among species. To change and adapt a body shape to the environment, the biological functions of the products of the morphogenetic genes must have been modified in each species (Holland, 2000, Chuong *et al*., 2000, Manzanares *et al*., 2000, Graham, 2000). We have compared the protein structures of the products of two morphogenetic genes related to the early development of medaka and zebrafish.

The first forkhead box gene was found in *Drosophila* by analysis of the mutant. The mutant analyzed exhibited irregular invagination of the foregut and hindgut at the larval stage (Jürgens and Weigel, 1988, Weigel *et al*., 1989). The gene encodes a transcription factor which has a unique DNA-binding domain called the winged-helix motif or forkhead box (Fox). In the Fox gene family, this encoded domain has a highly conserved sequence of 110 amino acids (Weigel and Jückle, 1990). Fox genes have been found in various organisms, from yeasts to humans (Zhu *et al*., 1993; Pati *et al*., 1997). In the early developmental stages of vertebrates, *HNF3β*, one of the Fox genes, is expressed in the node, notochord and floor plate of the neural tube. It has been suggested that the products of some of the Fox genes play a key role in tissue induction and organogenesis (Strähle *et al*., 1993). Based on the conserved DNA-binding domain, the Fox genes have been grouped into ten different classes (I-X) (Kaufmann and Knöchel 1996). Regarding the chordates, the Fox nomenclature committee grouped the Fox genes to 16 subclasses (FoxA-Q) (Kaufmann and Knöchel, 1996, Kaestner *et al*., 2000).

To understand the relationship between inductive interaction and the roles of Fox gene products, we have cloned novel Fox genes regulating the early developmental stages, and attempted to clarify the precise biological roles of the Fox gene products. Odenthal and Nüsslein-Volhard (1998) cloned nine Fox genes from a zebrafish embryo during gastrulation and neurulation. We cloned two Fox genes from a medaka embryo. Based on DNA sequence homology and their expression patterns in embryos, one gene, *MeHNF3β*, was homologous to the zebrafish *axial/fkd1* (or mammals *HNF3β*, FoxA2). The product of the other gene, *Mefkh1,* had a DNA-binding domain of the FoxA3 type. The predicted amino acid sequence of the Fox domain of the *Mefkh1* product was most similar to that of the *fkd2/zffkh1* product (Dirksen and Jamrich, 1995, Odenthal and Nüsslein-Volhard, 1998). However, the expression patterns were different in some aspects. These findings suggest that the Fox gene products can vary in their biological roles. Some of them, such as the *MeHNF3β* product, have retained its very conserved biological functions, whereas the molecular structure of others, such as the *Mefkh1*

* Corresponding author: Tel. +81-596-58-6411;
FAX. +81-596-58-6413.
E-mail: hiokamot@affrc.go.jp

product, has been changed or differentiated so that the protein is specifically involved in organogenesis in the species evolution.

MATERIALS AND METHODS

Medaka strain

The medaka (*Oryzias latipes*) *bbRR*-strain was obtained from a local pet shop and maintained at 28°C under a 14 hr light / 10 hr dark regime at the National Research Institute of Aquaculture (Mie, Japan). The eggs were harvested at room temperature (24–26°C). The developmental stage of the embryo is described in terms of the stage number according to Iwamatsu (1993).

Cloning of medaka Fox genes

We amplified the forkhead box fragments from medaka genomic DNA using the polymerase chain reaction method with primer-1, AAG CCC CC (TA) TAC TC (CG) TAC AT and primer-2, GCG (GC) AT AGA GTT CTG CCA. After cloning the 150 bp fragments into pBluescript II (Stratagene), we used one of the clones as a probe to screen the medaka embryonic cDNA library. Prior to use, we determined that this probe fragment was not homologous to both *MeHNF3β* and *Mefkh1*. Poly-A RNA was isolated from 500 medaka embryos from the mid-blastula to the late-gastrula stage (total RNA extraction kit and mRNA purification kit, Pharmacia). The embryonic cDNA was synthesized using an oligo-dT primer (Timesaver cDNA synthesis kit, Pharmacia). After screening 500,000 clones from the medaka cDNA library, two types of Fox gene were identified by DNA sequencing of both strands (ABI PRISM, Applied Biosystems).

For *Mefkh1*, we determined the sequence of the 3'-flanking region up to the poly A signal using the 3'RACE method (Marathon cDNA amplification kit, Clontech)

Southern blotting and Northern blotting

Southern blotting and Northern blotting were conducted according to 'Current Protocols in Molecular Biology' (Wiley, USA). We used Hybond-N, -N+, and -NX (Amersham, UK) membranes for DNA or RNA blotting. We labeled the DNA probe using the *Bca*BEST™ labeling kit (Takara, Tokyo).

```
ATGTGTTTGCGCGCTGCTGCTCCTCTGTGCGTCACTTGCAGACCGAGGGCCGCGCGGCCACAGGAAGAAACGGGTCACCACGAGAAGAAG    90
TGAAGAGCAAAACAAGATCAAGGTTTCCAACCATGGATGACTAAAACTTTCATTGACTTACCTGCAGGAACACGAGACGCGTTTCCCCCC   180
TGAAAACAAACGGAACCAACTTTTTCTCTGGATTTTTTTTTCTCTGCTGAGGATTTTGGATTTTTAGTTGTTTTTTAAGACCCTCTCAAA   270
GATAAAGACGATTCTTCTCCACAGACTTTAGTTGGGGATGTTGAGCTCTGTGAAGATGGAAGCGCACGATATAACAGACTGGAATACAAC   360
                                M  L  S  S  V  K  M  E  A  H  D  I  T  D  W  N  T  T

TTTCTACAGCGAGGCCAGCGAGATGTATTCCTCACCCAGCGCCACCGGGCTGGGCTCCATGGGCTCCATCAACAGCTACATCAACCTGAA   450
F  Y  S  E  A  S  E  M  Y  S  S  P  S  A  T  G  L  G  S  M  G  S  I  N  S  Y  I  N  L  N

CGCCGCTGCCTCCCCCACCGCTATGAACATGCCCTACCCCAGCTCCAGCCTCAGCAGCTCCTCTCTGGCCCCCATGGGCTCAGGGCCAAA   540
 A  A  A  S  P  T  A  M  N  M  P  Y  P  S  S  S  L  S  S  S  S  L  A  P  M  G  S  G  P  N

CCACATGTCCCTCTCCCCCGTGGCCTCCTCGCTCAGCTCCGGCTCCCTGACTCAGCTGGGCCCCGCCGCTGCGGGCTCGCTGGGCCCTCT   630
 H  M  S  L  S  P  V  A  S  S  L  S  S  G  S  L  T  Q  L  G  P  A  A  A  G  S  L  G  P  L

GTCCCACTACCAAAACATGGGTCAGTCCATGAGCCAGCTGGGTTACCCGTCCACCGGGTCTCTGAGCCGCTCCAGCCCCAAGGAGATCCC   720
 S  H  Y  Q  N  M  G  Q  S  M  S  Q  L  G  Y  P  S  T  G  S  L  S  R  S  S  P  K  E  I  P

GCCGAAGCCCTACCGGCGCTCCCTGACCCACGCCAAGCCACCGTACTCCTACATCTCCCTCATCACCATGGCGATCCAGCAGAGTGGCAG   810
 P  K  P  Y  R  R  S  L  T  H  A  K  P  P  Y  S  Y  I  S  L  I  T  M  A  I  Q  Q  S  G  S

CAAGATGCTGACTCTGAACGAGATCTACCAGTGGATCATGGACCTGTTCCCCTACTACCGCGAGAATCAGCAGCGCTGGCAGAACTCCAT   900
 K  M  L  T  L  N  E  I  Y  Q  W  I  M  D  L  F  P  Y  Y  R  E  N  Q  Q  R  W  Q  N  S  I

CCGACACTCGCTCTCCTTCAACGACTGTTTCGTGAAAGTCGCTCGCTCGCCGGACAAGCCGGGCAAAGGCTCCTACTGGACTCTGCACCC   990
 R  H  S  L  S  F  N  D  C  F  V  K  V  A  R  S  P  D  K  P  G  K  G  S  Y  W  T  L  H  P

GCAGTCCGGCAACATGTTCGAGAACGGCTGCTACCTGAGGCGCCAGAAGCGCTTTAAGATCGAGGACAAGGCCTCCAAGAAGGGAAGCAA  1080
 Q  S  G  N  M  F  E  N  G  C  Y  L  R  R  Q  K  R  F  K  I  E  D  K  A  S  K  K  G  S  K

GAACCAGGAGGGGGGCTCCGGGAAAGGCAGCCACTCTGGAGACCACCTGGTGGAGAACCACAGCCCTGCCGGTGGGTCGGAGGGGCCTGA  1170
 N  Q  E  G  G  S  G  K  G  S  H  S  G  D  H  L  V  E  N  H  S  P  A  G  G  S  E  G  P  D

CTCCGCCCACTCAGACAACTCCCACCCGGGCTCGTCGGACGACCAGCCCCATCAGAGGAACAGCCTGGTTCCTCTGGACTGCCCTCAGCT  1260
 S  A  H  S  D  N  S  H  P  G  S  S  D  D  Q  P  H  Q  R  N  S  L  V  P  L  D  C  P  Q  L

GTCATCCTCGCACCTCCATAGCCCACCTGTCTCCCTGCCCCCCTCCTCCTCCTCCCTGTCCCTCTCGGGGACCTCCTCCTCAAACCCGCT  1350
S  S  S  H  L  H  S  P  P  V  S  L  P  P  S  S  S  S  L  S  L  S  G  T  S  S  S  N  P  L

CCTCCACCCCCAGTCCTTAGCCAGCAGCTCCCACCTCCTCCCCAGCCCCATGCAGCAGCACATGGACTTACAGAGCGACCTGAAGTCCCT  1440
 L  H  P  Q  S  L  A  S  S  S  H  L  L  P  S  P  M  Q  Q  H  M  D  L  Q  S  D  L  K  S  L

GGACCCCCACTACAACTTCAACCACCCCTTCTCCATCACCAACCTCATGTCCAACGAGCAGAAGATGGACCTCAAGTCCTATCAGGACCA  1530
 D  P  H  Y  N  F  N  H  P  F  S  I  T  N  L  M  S  N  E  Q  K  M  D  L  K  S  Y  Q  D  Q

GGTGATGGCCTACAACAGCTACACCGCCGGCTCCCCGGTGGGGGCCAAGCAGATCTATGACAGCCCTGGCCCCGCCGCCATGGACTCAGG  1620
 V  M  A  Y  N  S  Y  T  A  G  S  P  V  G  A  K  Q  I  Y  D  S  P  G  P  A  A  M  D  S  G

CGCTTACTACCAAACTCTGTACAGCCGCTCTGTGCTCAACGCCTCTTAAAGGGACTCACAGGCACGTACACACACACACACACACGTCTT  1710
 A  Y  Y  Q  T  L  Y  S  R  S  V  L  N  A  S  *

TTTCTGCTCTAAAATGGACCCCCCCTCCCCCACCTGCCTGGCTCACGTCCAGCGGCCTTCACAGGACTCATCTGAAATGGCGGCCGCGTT  1800
TTTTTACCTCCTGGTTTGGGGGCCGGCCGAACCGATGAAGAGGACTCCTCTGTGAGTGTCGGCCCCGTCCTGACTGTACAAAATGTAAAT  1890
ACAAGAACCGTGTGTGTGTGTGGTGGGGGGGGGGGGGGACGTTTGCATTTATTTATGAAGGGAATAAATTTTAATATATCTTATGAAGGG  1980
CTCTGTCCACCGGTTACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  2061
```

```
ACCGCTCTAGAACTAGTGGATCCCCCCGGGCTGCAGTAATCCGCGGCCGCTCGCTCACTGACAGACGCGTCGCGGCTGGCTGTGAGCCAC   90
ACACATGCGCCAGCCCTCGACAAACGGACACGCCGACCACCCAAAAACCCTCAGTGGTCCCCAGTCACTTGCTGTTAGAAAGACTTTTTG  180
GCTTTTCCTCTGCTGGGAGACGCTTCGTTCAAGTCCTCCATGATGCTTGGAGCAGTTAAAATGGAAGGACACGAACACACCGACTGGAGC  270
                                                  M  M  L  G  A  V  K  M  E  G  H  E  H  T  D  W  S

ACATACTATGGAGAGCCCGAGTGTTACACCTCTGTTGGAAACATGAACACCGGCCTGGGAATGAACTCTATGAACACCTACATGAGCATG  360
T  Y  Y  G  E  P  E  C  Y  T  S  V  G  N  M  N  T  G  L  G  M  N  S  M  N  T  Y  M  S  M

TCCGGCATGAGCACCACCGCAAACATGACGGCCAACTCCATGAACATGTCATACGTCAACACGGGCATGAGCCCCTCCATGACCGGCATG  450
S  G  M  S  T  T  A  N  M  T  A  N  S  M  N  M  S  Y  V  N  T  G  M  S  P  S  M  T  G  M

TCACCAGGCACCGGAGCCATGAACGGGATGGGAGCAGGCATGACGGCTATGAGCACAGCCCTGAGCCCCAGCATGAGTCCTATGACCGGA  540
S  P  G  T  G  A  M  N  G  M  G  A  G  M  T  A  M  S  T  A  L  S  P  S  M  S  P  M  T  G

CAGCCTGGGTCTATGAACGCCCTGACATCCTACACCAACATGAACGCCATGAGCCCCATTTACGGACAGTCGAACATCAACAGGTCCCGA  630
Q  P  G  S  M  N  A  L  T  S  Y  T  N  M  M  N  A  M  S  P  I  Y  G  Q  S  N  I  N  R  S  R

GATCCAAAGACGTACCGGCGGAGCTACACGCACGCCAAACCTCCGTACTCCTACATTTCTCTCATCACCATGGCCATCCAGCAGTCACCC  720
D  P  K  T  Y  R  R  S  Y  T  H  A  K  P  P  Y  S  Y  I  S  L  I  T  M  A  I  Q  Q  S  P

AGTAAGATGCTGACGCTGGCCGAGATCTACCAGTGGATAATGGACCTCTTCCCCTTTTACCGACAAAACCAGCAGCGCTGGCAGAACTCA  810
S  K  M  L  T  L  A  E  I  Y  Q  W  I  M  D  L  F  P  F  Y  R  Q  N  Q  Q  R  W  Q  N  S

ATCCGCCACTCTTTGTCATTCAACGACTGTTTCCTCAAAGTCCCCAGGTCGCCGGATAAACCCGGAAAGGGCTCCTTTTGGACTCTTCAC  900
I  R  H  S  L  S  F  N  D  C  F  L  K  V  P  R  S  P  D  K  P  G  K  G  S  F  W  T  L  H

CCGGACTCCGGCAACATGTTTGAGAACGGCTGCTACTTGAGGAGGCAAAAGCGCTTCAAATGCGAGAAGAAGATGTCCATGAAGGAGCCC  990
P  D  S  G  N  M  F  E  N  G  C  Y  L  R  R  Q  K  R  F  K  C  E  K  K  M  S  M  K  E  P

GGCCGCAAGGGTGGAGACGGCGGCTCTGCCAACAGCAGCTCGGACAGCTGCAACGGCAACGAGTCCCCGCACTCTAACTCCTCCTCCGGC 1080
G  R  K  G  G  D  G  G  S  A  N  S  S  S  D  S  C  N  G  N  E  S  P  H  S  N  S  S  S  G

GAGCACAAAAGGTCCCTGTCGGACATGAAGGGGAGCCAGGCCCTGAGCCCGGAGCACACCGCGCCCTCCCCGGTGTCGCAGGGGCAGCAC 1170
E  H  K  R  S  L  S  D  M  K  G  S  Q  A  L  S  P  E  H  T  A  P  S  P  V  S  Q  G  Q  H

CTCATGTCTCAGCATCACTCGGTCCTGGCGCACGAAGCCCACCTCAAGCCGGAGCACCACTACTCCTTCAACCACCCCTTCTCTATTAAT 1260
L  M  S  Q  H  H  S  V  L  A  H  E  A  H  L  K  P  E  H  H  Y  S  F  N  H  P  F  S  I  N

AATCTCATGTCATCGGAGCAACAGCACCATAAAATGGACCTTAAGACTTACGAGCAGGTGATGCACTACTCTGGATATGGCTCCCCCATG 1350
N  L  M  S  S  E  Q  Q  H  H  K  M  D  L  K  T  Y  E  Q  V  M  H  Y  S  G  Y  G  S  P  M

ACCGGAGCGCTTTCCATGGGCTCCATGGCGGGTAAAGCCGGCCTGGATTCTGCATCTATACCCGACACATCCTACTACCAAGGGGTCTAT 1440
T  G  A  L  S  M  G  S  M  A  G  K  A  G  L  D  S  A  S  I  P  D  T  S  Y  Y  Q  G  V  Y

TCCAGGCCAATCATGAACTCCTCGTAGGCTCCCTGTATGCAGACTGCCCCCCACCCCCATTCCAGTTTGTACATTTGTAAAAAACAAAAA 1530
S  R  P  I  M  N  S  S  *

CAAAAAAATATAGACTTTGTGTACTATATGTATTTGAGATATTTTTGACCTTTCTCACCGTTTTAGTTGTTGAGCTTGTTAGACTGTAAG 1620
GAGAATAAAGTGTTGCGCTTCAAAGATGGACCCCAAAGACACAAAAACTGCTGTAAGATATTCACGTTTGCATAAGTGGACACAGAAACC 1710
AGTTT
```

```
class                                1        10        20        30        40        50        60        70        80        90       100       110
Ia FoxA2    H   HNF-3β      :TYRRSYTHAKPPYSYISLITMAIQQSPNKMLTLSEIYQWIMDLFPFYRQNQQRWQNSIRHSLSFNDCFLKVPRSPDKPGKGSFWTLHPDSGNMFENGCYLRRQKRFKCEK
            R   HNF-3β      :.....................................................................................................
            Me  MeHNF-3β    :.....................S.....A.....................................................................
            Z   Axial/FKD1  :.....................S.....................................................................D.
Ib FoxA1    H   HNF-3α      :.FK...P.................RA.S.................Y.....................V..A.........Y.....................
            M   HNF-3α      :.FK...P.................A.S.................Y.....................V..A.........Y.....................
            Z   FKD7        :.F....P.................A.S.................Y.....................V..S.........Y.....................D.
Ic FoxA4    X   XFD-1'       :....N.S.................A....M..N......V....Y.....................I......E......Y.....E.....................R
            X   XFD-1/pintalavis :....N.S.................A....M..N......I....Y.....................V......E......Y.....E.....................R
            Ax  fork head    :A...N.S.................V.......M..N......T....Y......S.....................V......E......Y.A.....................DR
            Am  AmHNF-3-1    :A.....................S..S.....V..A....F......Y.....................V....T..R.....Y.....EA.....................
            Z   FKD4         :V...TLS.....C.....A.R...N...D..RQ...Y.....................VR......S.....Y.A.....M.....Q.
Id FoxA3    H   HNF-3γ       :G...PLA.................A.G.................Y..E.....................V..A.........Y.A...S.....................L.E
            R   HNF-3γ       :G...PL..................A.G.................Y..E.....................V..A.........Y.A...S.....................L.E
            Me  Mefkh1       :P....L.....................GS.....N...........Y..E.....................V..A.........Y.....Q.....................I.D
            Z   FKD2/Zffkh-1 :P....L.....................QS.....N...........Y..E.....................V..A.........Y.A...N.....................I.E
            S   salmo-fkh    :P....L.....................GS.....N...........Y..E.....................V..A.........Y.A...N.....................IDQ
II FoxB1    Z   FKD3         :PG.NT.SDQ.....TA....SC.E...P.....KF...R..Y..E.T......L..N.......I.I..R..Q.......A...SC.D.....SF...R....VMT
                                         *********  *  *    *   * ***  *   ** ** * * **** ** *******    * *  ***** * ***  * *****   ** ****
```

Fig. 1. The cDNA sequences and the comparison of the forkhead boxes of medaka Fox genes.
(A) The *Mefkh1* cDNA length is 2,061 bp long. The coding region is from positions #308 to #1,666, and the forkhead box is surmised to be between positions #728 and #1,057. (B) The *MeHNF3β* cDNA length is 1,715 bp long. The coding region is from positions #220 to #1,464, and the forkhead box is surmised to be between positions #640 and #969. The forkhead box is underlined. The transactivating domains are region-II (double underline), region-III (dot lined square) and region-IV (thick lined square). The serine clusters are indicated in thin lined square. The *in frame* stop codons are indicated in thick underline. (C) The forkhead box, specific DNA-binding domain, is comprised of 110 amino acids. The medaka Fox genes are classified in class-I or FoxA. The class-I has four subclass-a-d (FoxA1-4). Subclass specific-amino acids are indicated in *square*. Am; *Amphioxus*, Ax; axolotl, H; human, M; mouse, Me; medaka, R; rat, S; salmon, X; *Xenopus*, Z; zebrafish.

We extracted whole genomic DNA from adult medaka in using 4M urea TNES buffer. DNA extraction was performed according to the method of Nakayama (1995). The extracted DNA was digested with nine restriction enzymes (*Apa I, BamH I, EcoR I, Hae III, Hind III, Kpn I, Pst I, Sac I, Xho I*) at 37°C overnight. Southern hybridization was performed at 65°C, and the blotted membranes were washed at high stringency, twice with 2×SSC /1 % SDS and twice with 0.2×SSC / 0.1 % SDS, at 65°C for 15 min each time.

To obtain total RNA from each tissue, we dissected an adult medaka that had been paralyzed with FA100 (Tanabe, Osaka), and eluted the RNA from each tissue using Tri-sol (Gibco-BRL, USA). We also used Tri-sol to extract total RNA from medaka embryos at various development stages. Northern hybridization was performed at 42°C overnight. Northern blotted membranes were washed at high stringency using the same buffers and for the same number of times as in Southern hybridization.

We used the entire sequence of *MeHNF3β* as a template for DNA labeling. For *Mefkh1* labeling, we used positions #142-#771 of the sequence of *Mefkh1* as a template. The autoradiogram signals were detected with a BioMax film (Kodak, USA).

Whole-mount *in situ* hybridization

Whole-mount in situ hybridization was carried out essentially according to the protocol of Hauptmann and Gerster (1994). Prehybridization and hybridization were performed at 55°C. The digoxigenin(DIG)-UTP-labelled RNA probes (Boehringer Mannheim, Germany) were generated from deletion clones retaining the N-terminus coding region (D59 for *MeHNF3β*; positions #0-#572 and R17 for *Mefkh1*; positions #11-#471) of the medaka Fox genes as a template (Fig. 1). For synthesizing the RNA probes, we used a DIG RNA labeling kit (SP6/T7)(Boehringer Mannheim), and a DIG nucleic acid detection kit for its detection (Boehringer Mannheim). The RNA probes were incubated overnight with embryos from the blastula to the hatching stage. The hatching enzyme gene (choriolysin H, HCE) was provided by Dr. Inohaya (Tokyo Institute of Technology).

Zoo blotting

The medaka and zebrafish were purchased from a pet shop and bred in our laboratory. The rainbow trout (*Oncorhynchus mykiss*) were bred at our field station. Japanese frogs (*Hyla arborea japonica*) and turtles (*Clemmys japonica*) were obtained from the field. A chicken was purchased from a grocery store. Rockhopper penguin (*Eudyptes chrysocome*) and short-fin pilot whale (*Globicephala melaena*) DNA samples were provided by Dr. Kage (Mie University). Mouse DNA was provided by Dr. Kudo (Tokyo Institute of Technology). All DNA samples were digested with *Dra I* and applied to a gel at 5 μg/lane. DNA probes were synthesized using the 5'- specific region of each medaka Fox gene (see whole-mount *in situ* hybridization). The probes were designed so as to delete the forkhead-box-encoding region. Southern hybridization was conducted according to the method of Church and Gilbert (1984). Prehybridization was carried out for 1.5 hr at 68°C, and the hybridization was performed at 65°C overnight. The washing procedure was the same as that used in Sourthern and Northern hybridization.

RESULTS

Sequencing of medaka Fox genes

To clone the Fox genes, we screened about 500,000 medaka cDNA clones using the PCR-amplified DNA fragment. We identified two types of Fox gene among the seven positive clones. After DNA sequencing, the sequence of the probe fragment was not found in these two Fox genes and the probe fragment thus belonged to another Fox gene.

As a result of the DNA homology search, we found that one clone was a novel Fox gene, *Mefkh1*. The *Mefkh1* cDNA was 2,061 bp long and the predicted coding region was from positions #308 to #1,666 (DDBJ; accession No. AB001573) (Fig. 1A). The region encoding the forkhead box was surmised to be between positions #728 and #1,057 (underlined in Fig. 1A). The predicted molecular weight of this gene product was estimated to be 48,973 daltons and the protein consisted of 453 amino acids. The other gene, *MeHNF3β* (medaka-*HNF3β*), was 1,715 bp long which included an open reading frame from positions #220 to #1,464 (DDBJ; accession No. AB001572) (Fig. 1B). In this coding region, there is a conserved region encoding the forkhead box between positions #640 and #969 (underlined in Fig. 1B). The amino acid sequence deduced from the DNA sequence is 415 amino acids long and the molecular weight of the protein was estimated to be 45,720 daltons.

Sequence comparison of the forkhead box

The Fox genes have been classified into various types and identified in various organisms (Kaufmann and Knöchel, 1996, Odenthal and Nüsslein-Volhard, 1998). We compared the conserved DNA-binding domain, the forkhead box, of the medaka gene products with those of other species. *MeHNF3β* is classified into subclass-Ia or Fox A2, while *Mefkh1* is grouped into subclass-Id, FoxA3 (Fig. 1C). The product of *MeHNF3β* consists of subclass-Ia-specific amino acids (squared column in Fig. 1C) and this gene is most closely related to the zebrafish *axial/fkd1*. The *MeHNF3β* product showed 99% identical to the *axial/fkd1* product in the forkhead box region and 88 % identity in the full-length amino acid sequence. The *Mefkh1* product is composed of subclass-Id-specific amino acids, and exhibits the highest similarity to the zebrafish *fkd2/zffkh1* and Atlantic salmon *forkhead* products (Fig. 1C). The *Mefkh1* product is 95% identical to the *fkd2/zffkh1* product in the forkhead box region and 58% identity in the full-length predicted amino acid sequence.

Expression of *MeHNF3β* and *Mefkh1* in adults and embryos

We determined which adult tissues express the Fox genes. Northern blotting analysis was performed using total RNA from the various tissues. The results showed that *Mefkh1* was clearly expressed in the liver, gut and skin, but the expression levels were low compared to those of *MeHNF3β.* We did not find differences in the *Mefkh1* expression level between males and females (Fig. 2A). On Northern blots, the *Mefkh1* signal was observed at about 2.0 kb. *MeHNF3β* was expressed in the gills, liver and gut (Fig. 2B), and we also found a weak total RNA signal from the brain (data not shown). These signals appeared to be around 1.7 kb. No major difference in the expression levels was found between the tissues of males and females.

We also determined when these genes begin to be expressed in the embryo. The *Mefkh1* signal was detected from the gastrula to the vessel-forming completion stage (Fig. 2C). The autoradiograms were overexposed to intensify the weak

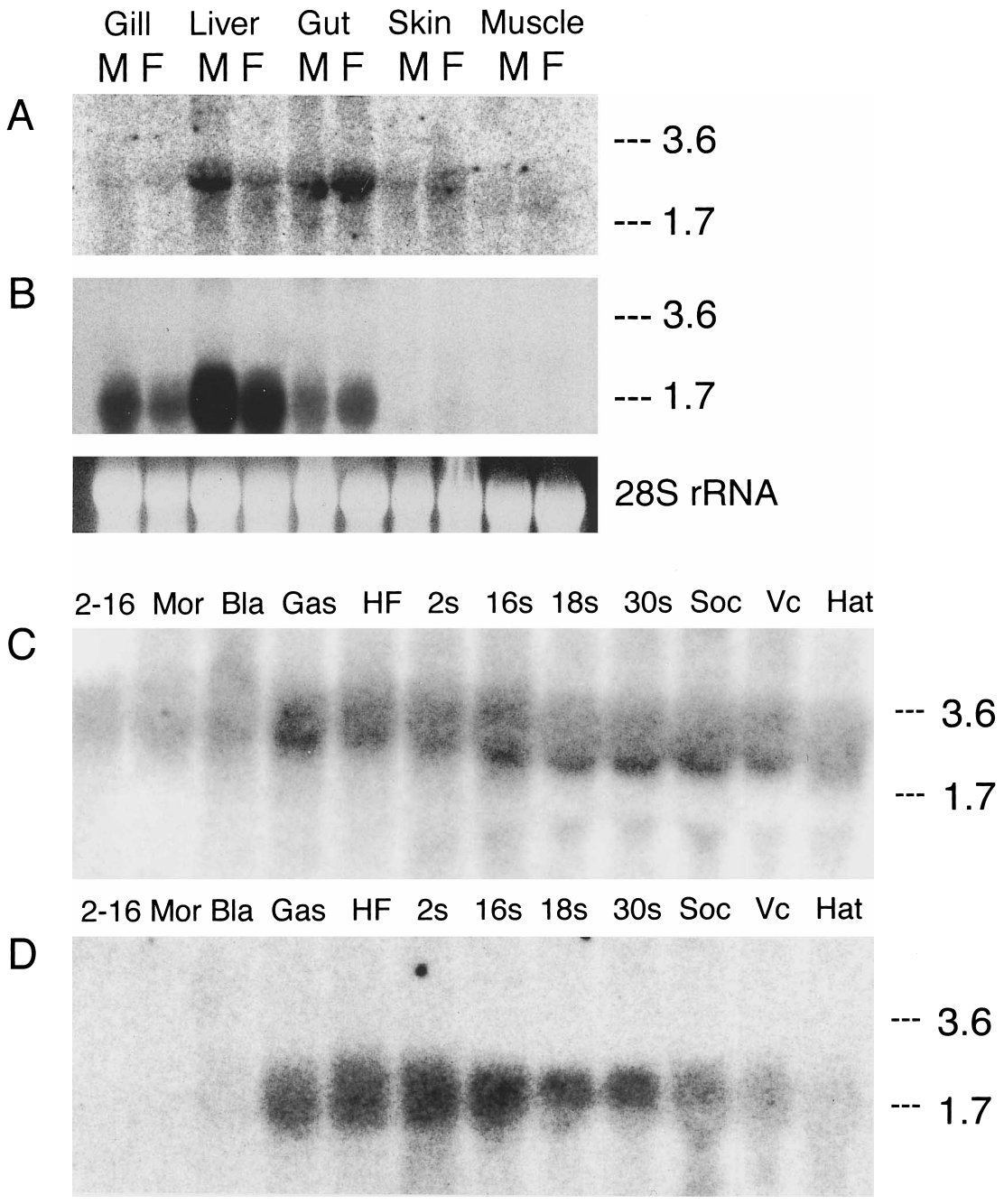

Fig. 2. Expression of *Mefkh1* and *MeHNF3β* in adults and embryos.
(A) *Mefkh1* probe hybridized with the total RNA that was extracted from liver, gut and skin. (B) *MeHNF3β* probe hybridized with the total RNA that was extracted from gill, liver and gut. Total RNA was applied (40 mg) in each lane. M; male, F; female. (C) The expression of *Mefkh1* in embryos. (D) The expression of *MeHNF3β* in embryos. 2-16; 2-16 cell stage, Mor; morula, Bla; blastula, Gas; gastrula, HF; head-forming stage, 2s; two-somite stage, 16s; 16-somite stage, 18s; 18-somite stage, 30s; 30-somite stage, Soc; somite-forming completion stage, Vc; vessels-forming completion stage and Hatch; hatching fry. Arrowheads indicate positions of transcripts of Fox genes. 28S; 28S ribosomal RNA (3.6kb), 18S; 18S ribosomal RNA (1.7kb). Each lane was applied at 30 μg of total RNA from whole embryos.

signals which were observed at around 2.0 kb. The *Mefkh1* expression peaked around the somite-forming completion stage. *Mefkh1* mRNA was not detected in the maternal unfertilized egg. The *MeHNF3β* signal was detected from the gastrula to the hatching stage (Fig. 2D). The maximum level of expression of *MeHNF3β* was observed around the two-somite stage to the 16-somite stage. This *MeHNF3β* expression peaked at an earlier stage compared with that of *Mefkh1*, and included the nine-somite stage at which the notochord tissue is visible (Iwamatsu, 1993). This implies that the *MeHNF3β* product plays a role in notochord development.

Whole-mount *in situ* hybridization of Fox genes

We analyzed the expression pattern of the medaka Fox

genes in embryos using whole-mount *in situ* hybridization. We used specific RNA probes to prevent cross hybridization among the Fox genes (*see* Materials and Methods).

Mefkh1 was expressed in the dorsal marginal zone and organizer at stages 12–13 (from the late blastula to early mid-gastrula stage). In stage-17 embryos (head-forming stage), a group of *Mefkh1*-positive cells appeared at the rostral end (Fig. 3A, arrow) which is the peak of a mass of cells known as the polster. This tissue includes precursor cells of hatching enzyme grand cells. *Mefkh1* expression was also observed along the body axis, but it was very weak. Notably, this *Mefkh1* expression pattern was different from that of the zebrafish *fkd2/zffkh1*. The axial tissue expression of *fkd2/zffkh1* was very similar to that of *MeHNF3β. MeHNF3β* expression was detected as a sharp stripe along the midline of the body (Fig. 3E, arrow). In stage-19 embryos (two-somite stage), the group of *Mefkh1*-positive cells was located at the tip of the head (Figs. 3B and 3C, arrow). At more advanced stages, the expression signal was not detectable in cells which were derived from the polster and which were moving into the inner pharyngeal duct. At stages 27–30 (24-35-somite stage), we also observed *Mefkh1* expression in the rudiments of the liver

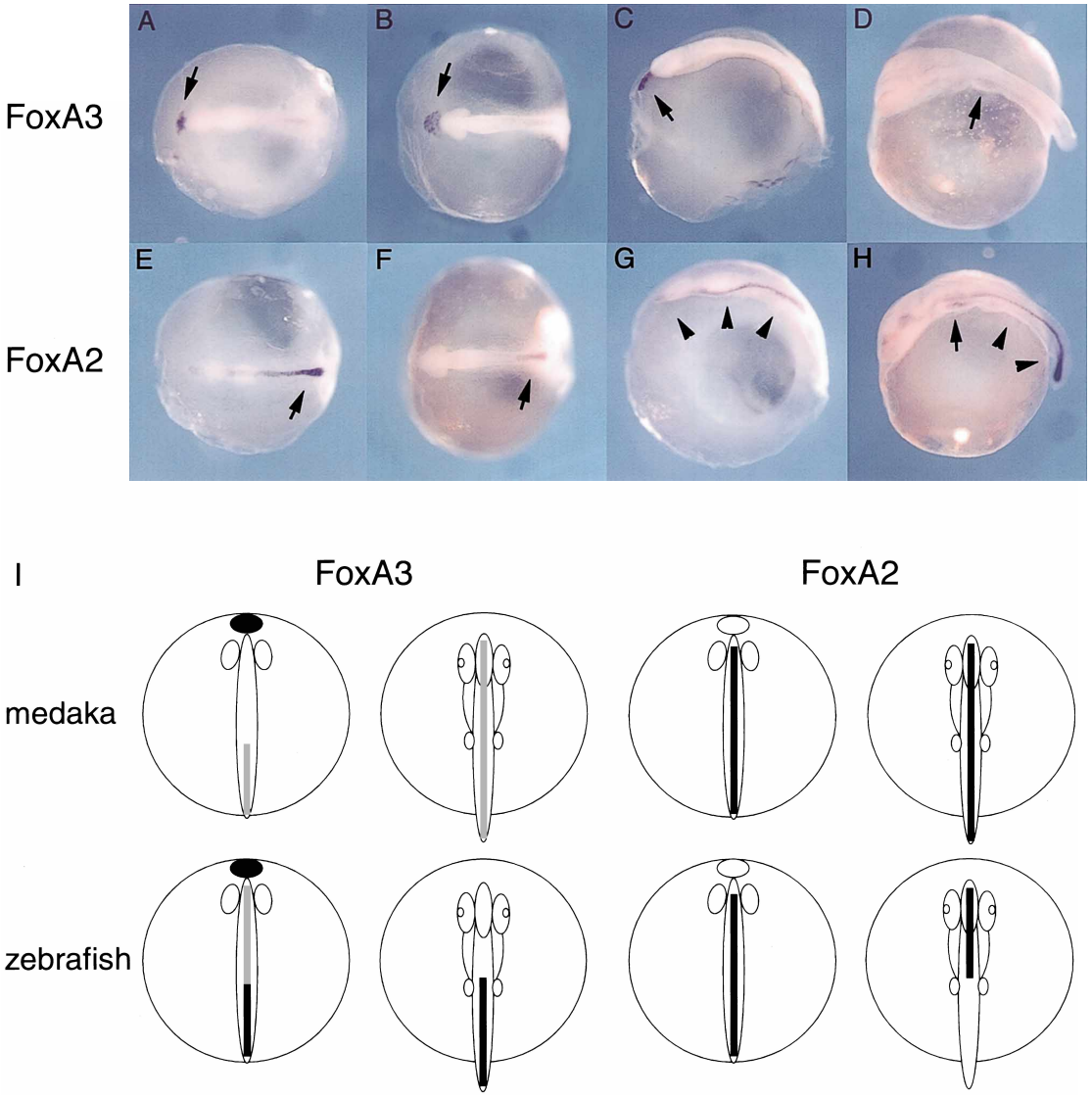

Fig. 3. Whole-mount *in situ* hybridization of *Mefkh1* and *MeHNF3β*.
(A) *Mefkh1*, at the head-forming stage, the expression was observed in the polster (arrow). (B, C) At two-somite stage, the polster cells expressed *Mefkh1* migrate into head region (arrow). (D) At 24-35-somite stage, *Mefkh1* was expressed in rudiments of liver (arrow). (E) *MeHNF3β* is expressed along body axis at head formation stage (dorsal view). (F, G) At two-somite stage, *MeHNF3β* was expressed in the primordium of notochord, ventral mesoderm (arrow and arrowhead). (H) At 24-35-somite stage, *MeHNF3β* was expressed in the otic vesicle (arrow) and the floor plate (arrowhead). Dorsal views (A, B, E, F) and lateral views (C, D, G, H), anterior to the left, posterior to the right. (I) The scheme of the expression pattern of FoxA2 and FoxA3. Both medaka and zebrafish FoxA2 express in axial mesoderm of their embryos. Whereas medaka FoxA3 restrictedly expresses in polster (and very slightly in posterior axial mesoderm), zebrafish FoxA3 expresses in polster and axial mesoderm like FoxA2 expression. Strong expression is indicated in *black*, and weak expression is indicated in *gray*.

(Fig. 3 D, arrow).

The expression pattern of *MeHNF3β* was very similar to that of mouse *HNF3β*. *MeHNF3β* was expressed in the organizer and dorsal marginal zone. In stage-17 embryos (head-forming stage), *MeHNF3β* expression was observed in the ventral mesoderm of the trunk and tail (Fig. 3E). This region is thought to be the primordium of the notochord. In stage-19 embryos (two-somite stage), the mRNA signal was also observed in the head region (Fig. 3G, arrowhead). At stages 27–30 (24-35-somite stage), the signals in the ventral mesoderm extended to the floor plate of the predicted mid-brain region and to the tail bud (Fig. 3H, arrowhead). We also observed signals in the otic vesicles (Fig. 3H, arrow). Our results indicate that *MeHNF3β* was expressed in the notochord and the precursor cells of the floor plate. At more advanced stages, we observed *MeHNF3β* expression in the primordium of the liver and a transient signal around the anus (stage 32, somite-forming completion stage). In *Caenorhabtis elegans*, the product of *lin-31*, a Fox gene, controls the fate of vulval cells (Miller *et al*., 1993). Thus, we speculate that the *MeHNF3β* product also plays a role in anus formation in vertebrates.

Southern blotting and zoo blotting

To determine the copy number of *Mefkh1* and *MeHNF3β* in haploids, we performed genomic Southern blotting using nine restriction enzymes (see Materials and Methods). *MeHNF3β* and *Mefkh1* were found to exist in a single copy in the haploid genome (data not shown).

To elucidate the conservation of the Fox genes among various vertebrate species, we performed zoo blotting analysis using medaka Fox gene probes. Hybridization of the *Mefkh1* probe with the digested genome was not clear for most of the species except turtle, chicken and penguin (Fig. 4A). The *Mefkh1* probe did not hybridize with genomic DNA of the zebrafish, *fkd2/zffkh-1,* which indicates that medaka FoxA3 has a molecular structure different from that of zebrafish FoxA3. We used the same digested genomic DNA on the agarose gels for zoo blotting of *Mefkh1* and *MeHNF3β.* The washing procedure for *Mefkh1* hybridization was carried out under high-stringency conditions as was that for *MeHNF3β* hybridization (*see* Materials and Methods). We concluded that the DNA samples were not degraded as shown in Fig. 4B. This result suggests that *Mefkh1* is conserved in all organisms, but the molecular structure of its protein product differs among species. Regarding the medaka genomic DNA, two bands were observed with estimated molecular sizes of approximately 850 bp and 6,500 bp. This indicates that *Mefkh1* was digested into two separate fragments, because the results of Southern blotting using the nine restriction enzymes showed that there is one copy in the genome. *MeHNF3β* is conserved in all species examined, from fish to mammals (Fig. 4B). This result shows that the *MeHNF3β* product has a universal biological function in all chordates and thus may be conserved in each species. For all of the DNA samples derived from the various animals used, one or two clear bands were observed except for the rainbow trout for which at least

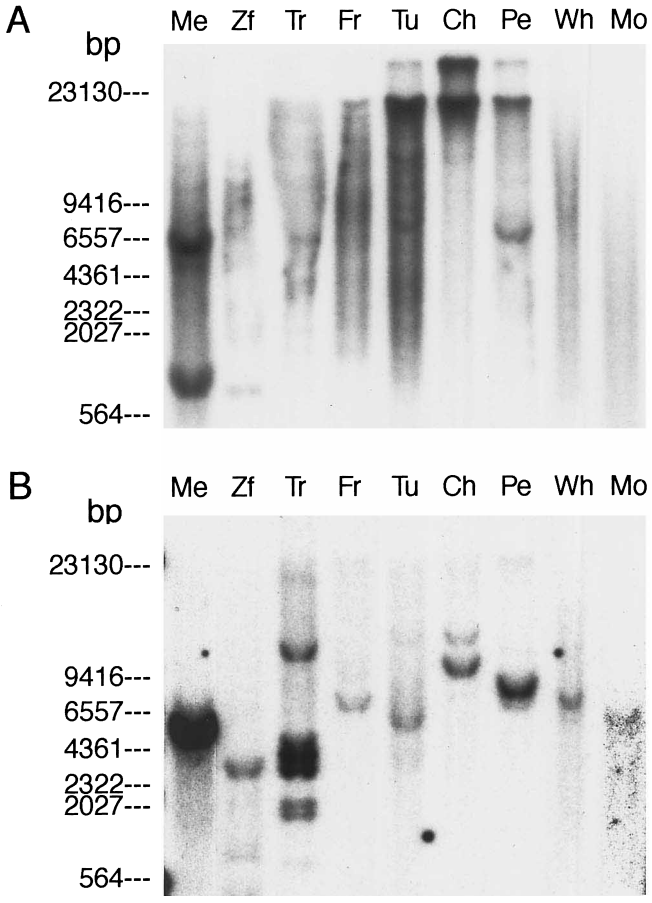

Fig. 4. Zoo blotting of *Mefkh1* (A) and *MeHNF3β*(B).
The medaka FoxA3 does not hybridize with the zebrafish genome DNA (A), although the medaka FoxA2 hybridizes with it (B).
Me; medaka, Zf; zebrafish, Tr; rainbow trout, Fr; frog, Tu; turtle, Ch; chicken, Pe; penguin, Wh; whale, Mo; mouse. Each lane was applied 5 μg of the genome DNA.

six bands were observed for the haplotype DNA. This result may indicate that rainbow trout may have originated from a tetraploid ancestor.

DISCUSSION

We cloned two medaka Fox genes both of which belonged to the FoxA type. FoxA2, the product of *MeHNF3β,* has a similar predicted protein structure and expression pattern in early developmental stages of embryos to those of zebrafish FoxA2. FoxA3, the product of *Mefkh1*, has a different serine cluster pattern and expression pattern different from those of zebrafish FoxA3. We would like to focus on the differences in FoxA3 characteristics between medaka and zebrafish.

Transcriptional activation domains and serine clusters

Analysis of the predicted protein structure revealed that the Fox proteins contain conserved motifs except for the DNA-binding domain. Regions II-IV are conserved in the class I, FoxA proteins (Pani *et al*., 1992). These regions are predicted to interact with transcriptional activation co-factors. The

MeHNF3β and *Mefkh1* products retain all of regions II-IV (Fig. 5). We found many serine clusters such as SS, SSS and SSSS in the Fox proteins. The serine clusters are mostly located outside of the DNA-binding domain and are marked on a schematic map of the Fox proteins (indicated by red bars) in Fig. 5. The *HNF3β* and *HNF3γ* proteins have highly similar serine cluster/scatter patterns between humans and rats (subclass-Ia and Id in Fig. 5). The medaka *MeHNF3β* and zebrafish *axial/fkd1* products appear also to have similar scatter patterns (subclass-Id in Fig. 5). This indicates that these are conserved proteins which have retained the same biological roles throughout species evolution. On the other hand, the serine cluster pattern of the *Mefkh1* product exhibits no similarity with that of the *fkd2/zffkh1* and Atlantic salmon *forkhead* products.

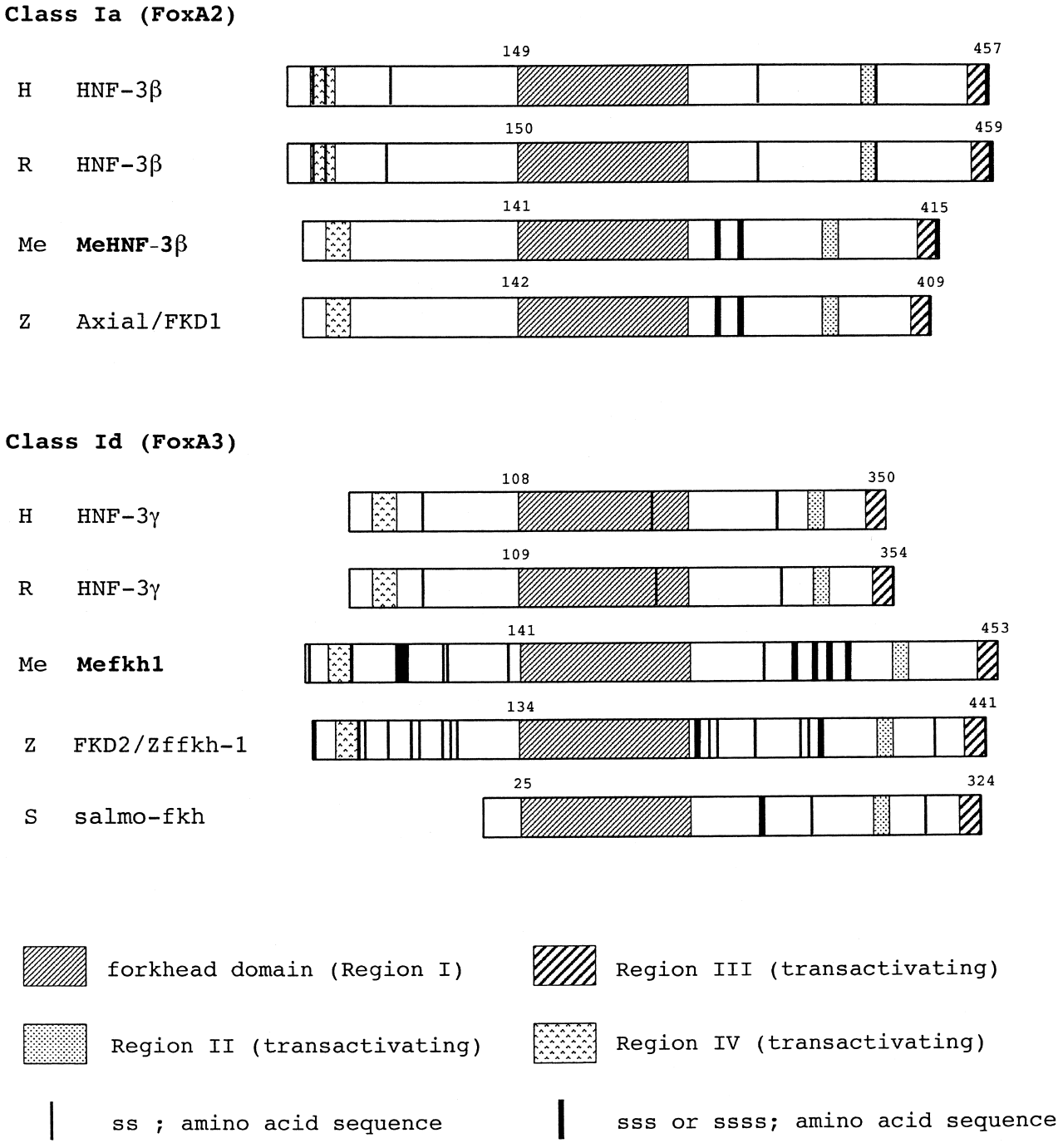

Fig. 5. Domain structures and the serine cluster index of the medaka Fox genes.
The DNA-binding domain is the forkhead box (narrow diagonal lined). The transactivating domains are region-II (dotted), region-III (wide diagonal lines) and region-IV (chevroned). The serine clusters are indicated in thin (SS) or thick (SSS or SSSS) bars.
H; human, Me; medaka, R; rat, S; salmon, Z; zebrafish.

This result suggests that *Mefkh1* is related to the zebrafish *fkd2/zffkh1* from the viewpoint of molecular evolution, but the biological roles of the respective products have possibly been modified during the long period of evolution after the organisms had diverged into different species. As a result of a Blast search, some of the serine residues in the cluster are considered to belong to the modification site. For example, Ser 269–271 and Ser 285–287 of the *MeHNF3β* product and Ser 299–300 of the *Mefkh1* product are included in the predicted phosphorylating site for casein kinase II. These serine residues possibly play a role in binding the transcriptional co-factor and regulating the gene expression.

Differences in expression patterns of Fox genes between two fish species

We compared the molecular structures and expression patterns of FoxA2 and FoxA3 between medaka and zebrafish. The medaka and zebrafish FoxA2 have almost the same molecular structures and expression patterns. On the other hand, the medaka and zebrafish FoxA3 have different molecular structures and expression patterns. Why are there differences in the degrees of conservation between FoxA2 and FoxA3? All class I (FoxA type proteins) genes in vertebrates are expressed in the derivatives of the embryonic shield. The *fkd2/zffkh1* gene (zebrafish FoxA3) is expressed in the embryonic shield, notochord, floor plate, polster (Dirksen and Jamrich, 1995) and neural crest cells (Odenthal and Nüsslein-Volhard, 1998). On the other hand, *Mefkh1* (medaka FoxA3) is expressed only in the embryonic shield and polster, and slightly in axial tissue (Fig. 3I). The expression pattern of *fkd2/zffkh1* in the notochord and floor plate is similar to that of the medaka FoxA2 gene. The biological function of the zebrafish FoxA3 might overlap with that of the medaka FoxA2, while the medaka FoxA3 function may have become specialized and differs from that of medaka FoxA2. Further comparative studies are required to elucidate the universal role of FoxA3.

Correlation between expression pattern of *Mefkh1* and hatching enzyme in polster

In fish, the polster is differentiated from the rostral end and derived from the embryonic shield (Shih and Fraser, 1996). Cells in the polster are assumed to be precursors of hatching enzyme (HCE) gland cells. As embryogenesis progresses, cells expressing the HCE in the polster migrate to the head region, and HCE expression is continuous until the cells come to reside in the inner wall of the pharyngeal duct (Inohaya *et al*., 1995, Inohaya *et al*., 1997). In stage-17 embryos, the group of *Mefkh1*-positive cells coincided with the cell mass expressing HCE. While HCE expression is continuously observed in the progenitor of the hatching enzyme gland proceeding to the inner wall of the pharyngeal duct, the *Mefkh1* expression level decreases as the cells started migrating to the head. This indicates that *Mefkh1* may be required for the hatching enzyme gland development.

Suggested role of *Mefkh1* product

Mefkh1 is transiently expressed immediately before the start of organogenesis. For example, its expression in the shield occurs before that in the epiboly, its expression in the polster occurs immediately before the development of the hatching enzyme gland, and its expression in the primordium of the liver occurs before development of the liver. This implies that the *Mefkh1* product may play a key role in cell differentiation and organogenesis.

ACKNOWLEDGEMENTS

We thank Dr. A. Kudo and Dr. K. Inohaya (Tokyo Institute of Technology) for kindly providing us with the mouse DNA and the choriolysin H (HCE) probe. We also thank Dr. T. Kage of Mie University for his kind gift of whale DNA. This study was supported in part by a grant-in-aid (Bio-design program) from the Ministry of Agriculture, Forestry and Fisheries of Japan (BMP-01-IV-1-2).

REFERENCES

Chuong C -M, Chodankar R, Widelitz RB, Jiang T.-X (2000) Evo-Devo of feathers and scales: building complex epithelial appendages. Curr Opin Genet Dev 2000 10: 449–456

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci U S A 81: 1991–1995

Dirksen ML, Jamrich M (1995) Differential expression of fork head genes during early *Xenopus* and zebrafish development. Dev Genet 17: 107–116

Graham A (2000) The evolution of the vertebrates – genes and development. Curr Opin Genet Dev 2000, 10: 624–628

Hauptmann G, Gerster T (1994) Two-color whole-mount *in situ* hybridization to vertebrate and *Drosophila* embryos. Trends Genet 10: 266

Holland LZ (2000) Body-plan evolution in the Bilateria: early anteroposterior patterning and the deuterostome-protostome dichotomy. Curr Opin Genet Dev 2000, 10: 434–442

Inohaya K, Yasumasu S, Ishimaru M, Ohyama A, Iuchi I, Yamagami K (1995) Temporal and spatial patterns of gene expression for the hatching enzyme in the teleost embryo, *Oryzias latipes*. Dev Biol 171: 374–385

Inohaya K, Yasumasu S, Araki K, Naruse K, Yamazaki K, Yasumasu I, Iuchi I, Yamagami K (1997) Species-dependent migration of fish hatching gland cells that commonly express astacin-like proteases in common. Develop Growth Differ 39: 191–198

Iwamatsu T (1993) The biology of the medaka (Japanese). Scientist co. Tokyo Japan 214–238

Jürgens G, Weigel D (1988) Terminal versus segmental development in the *Drosophila* embryo : the role of the homeotic gene *fork head*. Roux's Arch Dev Biol 197: 345–354

Kaufmann E, Knöchel W (1996) Five years of the wings of *fork head*. Mech Dev 57: 3–20

Manzanares M, WadaH, ItasakiN, TrainorPA, Krumlauf R, Holland PWH (2000) Conservation and elaboration of hox gene regulation during evolution of the vertebrate head. Nature 408: 854–857

Miller LM, Gallegos ME, Morisseau BA, Kim SK (1993) *lin-31*, a *Caenorhabditis elegans HNF-3/fork head* transcription factor homolog, specifies three alternative cell fates in vulval development. Genes Dev 7: 933–947

Nakayama I (1995) Application of genomic DNA analysis on fisheries science. (Japanese). Bulletin of Natl Res Inst of Aquaculture 24: 1–15

Odenthal J, Nüsslein-Volhard C (1998) *fork head* domain genes in

zebrafish. Dev Genes Evol 208: 245–258
Pati D, Keller C, Groudine M, Plon SE (1997). Reconstitution of a MEC1-independent checkpoint in yeast by expression of a novel human *fork head* cDNA. Mol Cell Biol 17: 3037–3046
Shih J, Fraser SE (1996) Characterizing the zebrafish organizer: microsurgical analysis at the early-shield stage. Development 122: 1313–1322
Strähle U, Blader P, Henrique D, Ingham PW (1993) *Axial*, a zebrafish gene expressed along the developing body axis, shows altered expression in *cyclops* mutant embryos. Genes Dev 7: 1436–1446
Weigel D, Jäckle H (1990) The *fork head* domain: a novel DNA binding motif of eukaryotic transcription factors? [letter]. Cell 63: 455–456
Weigel D, Jürgens G, Kuttner F, Seifert E, Jackle H (1989) The homeotic gene *fork head* encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo. Cell 57: 645–658
Zhu G, Muller EG, Amacher SL, Northrop JL, Davis TN (1993) A dosage-dependent suppressor of a temperature sensitive calmodulin mutant encodes a protein related to the *fork head* family of DNA-binding proteins. Mol Cell Biol 13: 1779–1787

(Received April 3, 2001 / Accepted May 23, 2001)