Gene encoding prolactin of red-spotted grouper, *Epinephelus akaara*, and its application as a molecular marker for grouper species identification

Bok-Ki Choi[1, #, †], Gyeong-Eon Noh[2, #], Yeo-Reum Kim[1], Jun-Hwan Byun[1], HanKyu Lim[3], Jong-Myoung Kim[1, *]

[1] Department of Fishery Biology, College of Fisheries Sciences, PuKyong National University, Busan 48513, Korea
[2] LED-Fishery Biology Convergence Production Research Center, PuKyong National University, Busan 48547, Korea
[3] Department of Marine & Fisheries Resources, MokPo National University, Muan 58554, Korea

Abstract

Groupers are economically important species in the fishery and aquaculture industries in Asian countries. Various species of grouper, including hybrids, have been brought to market even without precise species identification. In this study, we analyzed the structure and expression profile of the gene encoding prolactin (PRL) in the red-spotted grouper *Epinephelus akaara* based on genomic DNA and cDNA templates. The results showed that the PRL gene consists of five exons encoding an open reading frame of 212 amino acids, including a putative signal peptide of 24 amino acids and a mature structural protein of 188 amino acids. It showed amino acid identities of 99% with *Epinephelus coioides*, 83% with *Amphiprion melanopus*, 82% with *Acanthopagrus schlegelii*, 75% with *Oreochromis niloticus*, 70% with *Coregonus autumnalis*, and 67% with *Oncorhynchus mykiss*, indicating its closer similarity to *E. coioides* and other groupers but marked distinction from non-teleost PRLs. PRL mRNA expression was detected mostly in the brain, including the pituitary gland, with little expression in other tissues. While the 5-exon structure of the PRL gene of red-spotted grouper and the exon sizes were conserved, the sizes of the introns, particularly the first intron, were markedly different among the grouper species. To examine whether these differences can be used to distinguish groupers of similar phenotypes, exon-primed intron-crossing analysis was carried out for various commercially important grouper species. The results showed clear differences in size of the amplified fragment encompassing the first intron of the PRL gene, indicating that this method could be used to develop species-specific markers capable of discriminating between grouper species and their hybrids at the molecular level.

Keywords: Prolactin, Grouper, Hybrid, Molecular marker

Introduction

The aquaculture industry is continuously expanding its horizons to meet the demands for products. Various species of fish have been used in aquaculture to fulfill demands for products in various countries. The Serranidae (Perciformes) consisting of five subfamilies and including approximately 522 species belonging to 73 genera are fish of interest to aquaculture. The genus *Epinephelus* includes 89 species that are recognized as economically important marine fish, more than one third of which are cultured in Asia for domestic consumption and overseas export (Eschmeyer et al., 2010; Mai et al., 2014).

There have been a number of attempts to culture groupers from seed production to mariculture to overcome difficulties associated with insufficient seed supply. In Korea, the kelp grouper (*Epinephelus bruneus*) and red-spotted grouper (*E. akaara*) have been successfully induced to mature and produce fertilized eggs (Kang et al., 2020). However, the long time required to grow to marketable size in a temperate region represents an obstacle to their feasible application in aquaculture. There have been attempts to resolve these issues by interspecies hybridization with giant grouper (*E. lanceolatus*), which is one of the largest species of Serranidae with a rapid growth rate. Advances in hybridization of groupers resulting in the production of new species with intermediate phenotypes have led to concerns about the correct identification of *Epinephelus* species, which have often been misidentified in aquaculture (Noh, 2020; Rimmer & Glamuzina, 2019).

The red-spotted grouper (*E. akaara*), also commonly known as Hong Kong grouper, is one of the most commercially valuable marine fish for aquaculture in a number of Asian countries, including China, Taiwan, and Korea (Lee et al., 2020), despite its listing as endangered in the International Union for Conservation of Nature (IUCN) Red List of Threatened Species (Sadovy et al., 2018). The situation has become more complicated as the most expensive grouper species, including kelp grouper (*E. bruneus*), red-spotted grouper (*E. akaara*), and seven-band grouper (*E. septemfasciatus*), have often been misidentified. To overcome such difficulties, which were exacerbated by the generation of hybrid groupers with intermediate phenotypes, it is necessary to develop a rapid method for precise species identification.

A number of advances have been made in the development of molecular markers that are useful for species identification of marine organisms, including catfish, tilapia, salmon, oysters, and shrimp. Two PCR-based genetic analysis methods have been used for species identification, one of which involves amplification-based detection of the target organism with species-specific primers and the other employs amplification of fragments containing variable sequences, such as introns and microsatellites flanked by conserved regions with universal primers (Hebert et al., 2003; Monaghan et al., 2005). In these methods, species can be distinguished by analysis of the sequence or size of the amplified fragments. One strategy to sequence introns is to design primers corresponding to adjacent exon regions and amplify across the intron, i.e., so-called exon-primed intron-crossing (EPIC) markers (Li et al., 2010).

Prolactin (PRL) is a member of the protein hormone family and is involved in diverse biological activities, including water and electrolyte balance, growth, development, reproduction, immunoregulation, among many others (Manzon, 2002). In most euryhaline fish, it is generally accepted that PRL plays a key role in osmoregulation by controlling ion and water balance (Astola et al., 2003; Breves et al., 2020). While a great deal is known about the gene structure and amino acid (aa) sequence of PRL for fish species of interest to aquaculture, relatively little information is available specifically for groupers (Zhang et al., 2004). In this study, we analyzed the structure of the gene encoding PRL from *E. akaara* and propose the first intron of PRL as an EPIC molecular marker for distinguishing a major grouper species.

Materials and Methods

Genomic DNA extraction

Groupers belonging to six species were obtained from Chungsol Susan (Muan, Korea): red-spotted grouper (*E. akaara*, RG), kelp grouper (*E. bruneus*, KG), giant grouper (*E. lanceolatus*, GG), seven-band grouper (*E. septemfasciatus*, SG), and two hybrid species, i.e., giant-red (hGR) and giant-kelp (hGK). Tissues, including the brain with pituitary gland, gill, stomach, intestine, liver, kidney, and muscle, were dissected from RG following anesthesia with 3-amino benzoic acid ethyl ester at a concentration of 150 ppm (Sigma-Aldrich, St. Louis, MO, USA) for 1 min. All samples were frozen in liquid nitrogen and stored at –80℃ until use. Genomic DNA was isolated from the gill via phenol/chloroform extraction. Samples of approximately 0.1 g tissue were mixed with 630 μL TNE urea in microcentrifuge tubes containing 20 μL proteinase K and 70 μL 10% sodium dodecyl sulfate followed by incubation at 60℃ for 2 h. To the mixture was added 700 μL PCI (phenol:chloloform:isoamyl

alcohol = 25:24:1), which was mixed immediately by inverting and then centrifuged at 18,000×g for 5 min. The aqueous phase was mixed with an equal volume of isopropanol. DNA was precipitated by centrifugation at 18,000×g for 5 min, and the DNA pellet was washed with 75% ethanol. The pellet was suspended in 100 μL DNase-free water containing RNase A and incubated at 37℃ for 30 min. DNA was stored at –20℃ until analysis.

PCR amplification of the gene encoding PRL and cloning

PCR primers (Table 1) were designed based on the conserved region of PRL genes of several fish on GenBank site (Noh et al., 2013b). PCR was performed in 20 μL 1 × HiQ-PCR mixture (GenoTech, Daejeon, Korea) containing 10 mM Tris (pH 9.0), 1.5 mM MgCl₂, 40 mM KCl, 1 mM dNTPs with 0.2 μg genomic DNA of RG and 1 μM primer pair (F/R). Each reaction was carried out with an initial denaturation step at 94℃ for 4 min, 30 cycles of reactions composed of denaturation at 94℃ for 30 s, annealing at 60℃ for 30 s, and extension at 72℃ for 2 min, followed by a final extension step at 72℃ for 5 min. Amplified PCR products were analyzed by 1% agarose gel electrophoresis followed by staining with ethidium bromide (0.5 μg/mL). PCR products ligated into an All-in-one™ PCR Cloning Kit (Biofact, Daejeon, Korea) were transformed into *E. coli* DH5α as described previously (Inoue et al., 1990). Plasmid DNA was isolated using a plasmid purification mini kit (Thermo Fisher Scientific, Seoul, Korea) according to the manufacturer's proto-

col. Sequencing was carried out by GenoTech.

Sequence analysis

The full-length aa sequences of PRL were aligned with other PRL genes obtained from the National Center for Biotechnology Information (NCBI) GenBank database. Phylogenetic trees were constructed using the neighbor-joining method (Saitou & Nei, 1987) with MEGA ver. 5.2. The reliability of each tree node was assessed by bootstrapping with 1,000 replicates. A number-of-differences model was set as the substitution model, and gapped or missing aa residues were treated by pairwise deletion. Aa sequences used to build the phylogenetic tree were obtained for *E. coioides* (AAR97730), *A. schlegelii* (AAX21764), *A. melanopus* (AEB00558), *O. niloticus* (NP_001266715), *C. autumnalis* (CAA80660), *E. lanceolatus* (XP_033500110), *Anguilla anguilla* (XP_035254403), and *O. mykiss* (NP_001118205). PRL sequences from *Papio hamadryas* (ADG56475) and *Homo sapiens* (AAH88370) were used as outgroups.

Reverse transcription (RT)-PCR analysis

To analyze the tissue-specific expression profile of PRL in red-spotted grouper, RNA was isolated from the brain, gill, stomach, intestine, liver, kidney, and muscle using RiboEX™ reagent (Cambio, Cambridge, UK) according to the manufacturer's protocol. Tissue samples (~0.1 g) homogenized in 1 mL RiboEX™ reagent (Cambio) were mixed with 200 μL chloroform followed by centrifugation at 12,000×g for 10 min at 4℃. The supernatant was mixed with an equal volume of isopropyl alcohol and centrifuged at 12,000×g for 5 min at 4℃. The RNA pellet was washed with 80% ethanol and then treated with RNase-free DNase I at 37℃ for 20 min followed by purification on a HybridR column (Tribioscience, Sunnyvale, CA, USA) according to the manufacturer's protocol. Quality of total RNA was analyzed via 1.8% agarose gel electrophoresis followed by staining with ethidium bromide and quantification using a microvolume nucleic acid spectrophotometer (ASP-2680, V4.1; ACTGene, Piscataway, NJ, USA).

First-strand cDNA was synthesized using an M-MLV cDNA synthesis kit (Enzynomics, Daejeon, Korea) according to the manufacturer's protocol. The reaction mixture containing 80 μM oligo dT primer and 1 μg total RNA was incubated at 70℃ for 5 min and chilled on ice for 5 min. Reverse transcription was performed by adding 20 μL of a mixture containing 10 × M-MLV RT buffer, 2 μM dNTP, 40 U RNase inhibitor, and 200 U M-MLV reverse transcriptase and incubating it at 42℃

Table 1. PCR primers

Primer	Sequence (5′–3′)
PRL-F1	AGGAAACAGCAACAAGCAGGTC
PRL-E1F	AGAGATGGCTCAAAGACAC
PRL-E2F	CAGACTGCACTCCCTCAGCA
PRL-E3F (qRT-F)	TGACAAGGAACAAGCTCTACAAG
PRL-E4F	CTCTGCTAACACCCTGCCTC
PRL-E2R	CGTGGCTGAGAGTCGTGC
PRL-E3R	CTTGTCACTGGGTGTCTGCAG
PRL-E4R (qRT-R)	GAGGCAGGGTGTTAGCAGAG
PRL-E5R	AGGACTTTCAGGAAGCTGTC
PRL-R1	ACTCCTCCTCATCTTCACCTC
RT-F	CTCTGCTAACACCCTGCCTC
RT-R	GAAGCAGGACAACAGGAAGTT
β-Actin-F	AGAGATGGCTCAAAGACAC
β-Actin-R	CTTGTCACTGGGTGTCTGCAG

PRL, prolactin; F, forward; R, reverse; RT, reverse.

for 60 min. The reaction was stopped by inactivation at 72℃ for 10 min and then the sample was stored at –20℃ until use. PCR was carried out to amplify the PRL gene using the primers listed in Table 1 together with a primer set for β-actin from olive flounder (*Paralichthys olivaceus*) as an internal positive control. The specific primer assays were carried out with pituitary cDNA serial dilutions, showing amplification efficiencies approaching 100%.

Analysis of the first intron-crossing region of the PRL gene

PCR was carried out with primers E1F and E2R designed corresponding to the first and second exons of the PRL gene, respectively, and genomic DNA as the template according to the procedure outlined above. The sizes of PCR products were analyzed by 1% agarose gel electrophoresis followed by staining with ethidium bromide.

Low salinity acclimation experiment

Sixty red-spotted grouper (8.2 ± 0.1 cm total length, 8.3 ± 0.4 g body weight) were divided into 10 fish in six tanks (0, 4, 8, 24, 48 and 144 h). All fish were adapted to each tank containing seawater (32 psu) before the experiment. With the except of the control group (0 h), the salinity of the experimental tanks was decreased to 8 psu by adding freshwater. Ten fish were sampled at the different exposure times of 0, 4, 8, 24, 48 and 144 h. Fish were dissected after anesthesia by 150 ppm 3-aminobenzoic acid ethyl ester (Sigma-Aldrich) for 1 min to collect brain, gill, stomach, intestine, liver, kidney, and muscle. Collected samples were frozen in liquid nitrogen and stored at –80℃ until use.

Results and Discussion

Structure of the PRL gene of red-spotted grouper *E. akaara*

Groupers, including the red-spotted grouper *E. akaara*, have been of interest for their superior market value as well as growth rate depending on the species. To explore their potential for development as a successful aquaculture species, it is important to understand the characteristics of the key molecules, including PRL, involved in various aspects of their physiology affecting the efficiency of the aquaculture system. To examine the primary structure of PRL in red-spotted grouper, the full-length gene was obtained by PCR amplification and DNA walking of genomic DNA and cDNA templates. The results showed that the gene encoding PRL consisted of five exons of 42, 114, 108, 183, and 189 bp in order from 5′ to 3′, yielding an open read-

ing frame encoding a protein of 212 aa residues composed of a putative signal peptide of 24 aa and a mature protein of 188 aa. The gene included four introns of 1,293, 183, 1,908, and 1,035 bp in order from 5′ to 3′, and all sequences found at the borders of introns contained the consensus GT and AG splicing signals (Fig. 1).

The aa sequence of PRL was deduced from its nucleotide sequence. Comparison of the aa sequence of RG PRL with those of other teleosts was performed using ClustalW and the results indicated an identity of 99% with *E. coioides*, 83% with *A. melanopus*, 82% with *A. schlegelii*, 75% with *O. niloticus*, 70% with *C. autumnalis*, and 67% with *O. mykiss*. RG PRL also contains four cysteine residues in the loci conserved in the PRLs of other teleosts (Fig. 2). Phylogenetic analysis indicated similarity between *E. akaara* PRL and those of other teleosts but a broad distinction from non-teleost PRLs (Fig. 3).

Structural characteristics of PRL in fish

PRL is a multifunctional hormone found in all vertebrates together with PRL-like protein in jawless fish, one of the oldest groups of living vertebrates (Gong et al., 2022). Despite its wide range of functions, all PRL genes share similar structural motifs, including five exons interrupted by four introns with an exceptional exon 1a found in humans. We found that a single form of PRL with a putative signal peptide of 24 aa and a mature protein of 188 aa shared 99% aa sequence identity with that of orange-spotted grouper *E. coioides*. These structural characteristics are consistent with PRLs reported from other fish, including salmon, tilapia, sea bream, and other groupers (Astola et al., 2003; Swennen et al., 1992; Xiong et al., 1992; Zhang et al., 2004). Overall, teleost PRL was synthesized as a prohormone with a signal peptide of 23–24 aa and full-length aa sequence of varying length depending on fish species (Manzon, 2002): 185 aa in *A. anguilla* (Quérat et al., 1994), 186 aa in *Cyprinus carpio* (Chao et al., 1988), 187 aa in *Hypophthalmichthys molitrix* (Chang et al., 1992), two PRL forms of 188 aa and 177 aa in *Oreochromis mossambicus* (Yamaguchi et al., 1988), *O. niloticus* (Rentier-Delrue et al., 1989), and *Oncorhynchus keta* (Song et al., 1988) in contrast to ovine and human PRLs with 199 aa residues (Cooke et al., 1981; Verma et al., 1989). The presence of four conserved cysteines that may lead to the formation of two disulfide bridges in the C-terminal region is consistent with most teleosts except sturgeon and lungfish, which have three disulfide bonds, similar to mice and humans. Nile tilapia PRL188, which lacks the N-terminal disulfide bridge, appears to be more

```
ATGGCTCAAAGACACAGCGATGGAAACAAACTCCTCATGACGgtgagatgagctcagcttgtgtgtttatcttct   76
 M   A   Q   R   H   S   D   G   N   K   L   L   M   T                          14
tcttctgtaaagtcatctgctgcagtgaagaacctcacgtgaagtataaggttggatgatcatgaagaatttgagttgtgtgtctcagaggggaaga   172
gctgggcctgagatagaaggcagaggattctgtttgaggactgagcaggaagacctgatcacacagacagtgtttcagcagctggaggcagctttt   268
aatgagactgaagcttttttgattgtgtgtttgtgttgctatgcgcctcgagtttctgcgctctgaagtccttgagttgtaagaaaaaatcctctca   364
aagcggaacaacgccccacgtcatctgttctttcgtaatctttttgccccagtgttcaatgggatgatttgagaggcggggcttctgtggtggtca   460
cgacaacaggtttacagttggtaaacaatggaggagaaactggtggtagtggctgctggatacccagagctatacgactttacaaaccacagctag   556
caccatctgaatagaaaacagcagataagtgcagtactgtaaacgtacatgatgggagtggttggaagaagactcttcttaagcctgaaaatattga   652
aatgaaatgtaaatataagcacattggaagtcaacattcattaaagtaacatatctttcatttcttttttccaaatgacctgggaag   748
ctgatgagcaaacagcctcttaaaaatggattattaagtgttatatcgtgcacaacttgctcaagcaggtccagcggcctacgatatagcacttta   844
gattaccatgacctggatgactaagaatcttcaccaacgattcagatgcaaggtcacactgaccttgacctttgaccaccagaatctaatcagttc   940
attgtgaagtccaagtggacagttgcaccaaatttgaaaaaatattccctgagggtattcttaagatgtcgcgttcacaagaatgagatggatgca   1036
aggtcacagtgaccttaacctttgaatcttgaccaccaaaatctaattaggtcatcctcaagtgcaaataagagttcgcgccacatatgaagatat   1132
tccttcaagacattcttgagaatgaaacagacaaggtcacagtgaccttgacctttgatctttgaccactaaaaacgaattagtttacagttgaat   1228
tcaaggtgacatttgttccaaattaaaataaattcccttgacatattgtgttaacataacaacccctctgaatgttctcgtcaccgtcctgtgct   1324
ccacagGTGTTGTACATGGTGGCAGCGTGCAGCGCCGTCCCCATCAACGACCTGCTGGACCGGGCC   1390
      V   L   Y   M   V   A   A   C   S   A   V   P   I   N   D   L   L   D   R   A   34
TCTCAGCGCTCCGACAGACTGCACTCCCTCAGCACGACTCTCAGCCACGACCTGgtcagtcctgattgtt   1460
 S   Q   R   S   D   R   L   H   S   L   S   T   T   L   S   H   D   L                52
gtgggtgtcagtcctgattgttgtgggtatgaactgtgtcataccacgttatacaagtctggtttcttacgtattcagcacagcgagaatattcta   1556
caacatgacacacaaggtaatataataattcgtcctctgtccccctccagGACTTTAATTTCCCTCCTATAGGCCGGATG   1635
                                                  D   F   N   F   P   P   I   G   R   M   62
GCTATGCCCCGCCCCTCAATGTGCCACACCTCCTCTCTGCAGACACCCAGTGACAAGGAACAA   1696
 A   M   P   R   P   S   M   C   H   T   S   S   L   Q   T   P   S   D   K   E   Q   83
GCTCTACAAGTATCAGgtgagttactgttaccctttttatgtattccttaatctttcatttgtctcacattaccacgtcttaaccaagct   1787
 A   L   Q   V   S                                                                     88
Cgcagcaggcgttagagttggctccaccctgtgagttactgttacccttttatgtattccttaatctttcatttgtctcacattaccacgtctta   1883
accaagctagcagcaggcgttagagttggctccaccctcttgtcataatatggtcactactggccccaaaaatccgagatggcgacgttcaaaatg   1979
ccaaattcaaggtttcaaaatgggagtttgcaaaccaatgggcggcgtcatggtggttacattcgttatgttgtctttaaccaactcatctcactg   2075
tgacctcgtcacatgtccacgtttggtcatggacttccacgtccacatgtgacgtccaaggtaccctgggtgtgttggttgttacatgcattgtc   2171
tgttttcaaaatacactctgttttcacaggaaatgtacagtttgcatacagtctctcttcaaaataaaagcactacgtcggtacaacactgcaaat   2267
tggtgtttgttctattcaataaaacacgcacatggttacatttaggaaaaaagaacagtgtttggctttactatcgtacaggaagtgaacactggc   2363
ctcctgggagaaagtctgtggttgttggatccaccaccacccctcctgtccgtcctattcggacttctgcgtgccttaactctgttgttgtccca   2459
ccatatttgcccactgtttgctgtgaaacaaaaagggcaaccgactggctgcgtatcgtgccgatgtgagaggacaggtgttttttcgttgtctga   2555
cagcagaagtcactgaccaagcactggttctcgaagacttcagagtgagaccgggttgtgttgtctttttctctgtctctgttttacatcagatca   2651
gatcttctaaccagctggagtcaagtagtttgagctgaaatgtttccaaagttcaaattattcttattatagagagagaggagaaattaagcagac   2747
atccaaagattggagctctaaaacagacgaaaggacacaaaacagaaggcttctcaagactgtttcagtgattactattgattcattaacgtaaag   2843
ctcacactgactggcagcctggtggcttagtgggtgtacacagcctctgggtcaatgaatgctgggagtggtagtgccgagagtggactgatgatg   2939
atggctgataaaaatagtttaggatgcaattcactcatttcatttgaaggaatacagtttttaagtgacgagcactcaaatgttttttacacaaaat   3035
atgaaaccagcgccagcagccaacctccaggaagtcatcgccaccaagccaagaaatggtcgagcacattatgcctcgtaaaacaagtgttgtttt   3131
atactgaggtttttgtgcaaggtataacatgttaactagcgagctttacacctcgtttacacatagttgtgtgtacgtaagcgagtcaagcaggaa   3227
gtccattcattcctagatctccaatgtgtctgcaaaactcctccatctgttttgttttcggttgggaacttgcctttaaacacatcagaaagtttg   3323
aagctatagtacattgttgtcagaaaagcgaggtttgatattagccgagactttcccagaatgggcacaaacgaacatcactgagctgtctcagcttt   3419
cacccgaataactgaagcgatcattgtagatgatctaaagaggattcaagtacccatcattctgtcctaaatctcagcaatgtgctgccaaacaat   3515
gtgatgaaaaatctgtaagattgaaattataaatgtgtcattttatttctttattgtgacattacctctgcaggccgacagctggaattacgagtg   3611
gtggtcaatctatctgagcagactattgctattggtttctcacagcatttcatctcccgtctctccagttgtccttcagctcattgtctcatctca   3707
tgattctcctctcctccagGAGTCCGATCTGCTGTCATTGGCTCGCTCTCTGCTTCAGGCCTGGGCCGAC   3777
                   E   S   D   L   L   S   L   A   R   S   L   L   Q   A   W   A   D   105
CCCCTGGTCGTCCTGTCCACCTCTGCTAACACCCTGCCTCACCCGGCCCAGAGCACCATCTCC   3840
 P   L   V   V   L   S   T   S   A   N   T   L   P   H   P   A   Q   S   T   I   S   126
AACAAGATCCAGGAGCTGCAGGAGCACACCAAGAGCCTGGGAGACGGCCTGGATATCCTGTCT   3903
 N   K   I   Q   E   L   Q   E   H   T   K   S   L   G   D   G   L   D   I   L   S   147
GGAAAGgtgtaagcacaaaaagaccatagactgtatagaataacaacccggtctaactccagagtcatcaagaaccagcacttcgtcagac   3997
 G   K                                                                               149
actgacgaaaaaagccttcctttcggcgcccagtgacgtcaggggggaaccccggtgttgttcaaccaaacgctgaacgagactttctattccacc   4093
aaaacgttacattactggagcagagactggagcatgcgcagtcgacacgcttgctatagctacagtcgtgtcgagtcacaaatattgggctgcaagt   4189
ggccgtctgcaaagggggtccgtgcagaccctccgggggacatgagtggatcacgtgtgtatagtcacgtaactttcacgagcagaaaacagactcac   4285
aaagcgatagctccacgtactctgtggttacatcatcggtatgttgcttaaccaacgtggttgctcagacaggcattacattttagtcgggttttg   4381
ttctccataccttacgactaaaagtcatgattctacatcaaaatgacgtcagcatgagacgtttggaagacaatggattttgggtcacttaacaac   4477
acaatttaaaaccaacaaaatatcaacgtcagttttcttcaaattaacgttggcgttagatgtgcaattgaaattcacccagcgtcacaacctaaa   4573
ttcaagcaaatatagacgtgtaacaacattcgtggccagctatacaggtgtattcagcctctgtacagtcgtcataaatgttgaa   4669
attagctcagagaccaaactgcttttagtaccaggctgtaaacatgttcatttctgctgtaaagtcttaacatgggggtctgtggggactgactct   4765
tggagccagcctccagtggacattagaggaactgcagttttttgtcacttccatgttcagcctgtttcctgcttgcgacaacaagtccttgtttt   4861
tgtatataaatcggtactataattgaaaatgttaatataaaatgtatttttgagactgtttttcctccgcctctttttctctcccagATGGGT   4951
                                                                                     M   G   151
CCAGCGGCTCAGCTCATCTCCTCGCTGCCCTACAGAGGAGGCAACGACATCGGCCAGGACAGG   5014
 P   A   A   Q   L   I   S   S   L   P   Y   R   G   G   N   D   I   G   Q   D   R   172
ATCTCCAAATTGGTCAACTTCAACTTCCTGTTGTCCTGCTTCCGCCGCGACTCGCACAAGATC   5077
 I   S   K   L   V   N   F   N   F   L   L   S   C   F   R   R   D   S   H   K   I   193
GACAGCTTCCTGAAAGTCCTACGCTGCCGGGCAGCAAAACTGAGACCGGAGATGTGCTAA   5137
 D   S   F   L   K   V   L   R   C   R   A   A   K   L   R   P   E   M   C   -   212
```

Fig. 1. Sequence of the gene encoding prolactin of red-spotted grouper _Epinephelus akaara_. Nucleotides corresponding to the coding regions in the exons are shown in uppercase and noncoding regions are shown in lowercase. Amino acid sequences corresponding to exons are indicated in bold. Red circles indicate GT and AG sequences corresponding to splicing of the introns.

```
E. akaara        ---MAQRHSDGNKLLMTVL----Y----------MVAACSAVPINDLLDRASQRSDRLHS   43
E. coioides      ---MAQRHSDGNKLLMTVL----Y----------MVAACSAVPINDLLDRASQRSDTLHS   43
A. schlegelii    ---MARRETNGSKLFITVL----C----------MVAACGAVPISDLLDRASQRSDTLHS   43
B. melanopus     ---MAQRRTNGSKLLLTVL----Y----------VVAACKAVPISDLLDRASQRSDRLHS   43
O. niliticus     ---MRQRRTSGTNLFMTVL----C----------VVAMCRAVPINELFERASQHSDKLHS   43
C. autumnalis    ----MARRSQGTKLHLAVL----C----------LVVSCHAIGLSDLMERASQRSDKLHS   42
O. mykiss        ----MARRSQGTKLHLAVL----C----------LVVSCHAIGLSDLMERASQRSDKLHS   42
P. hamadryas     MNIKGLPWK-GSLLLLLVSNLLLCQNVAPLPVCPGGAARCQVTLRDLFDRAVVLSHYIHN   59
H. sapiens       MNIKGSPWKAGSLLLLLVSNLLLCQSVAPLPICPGGAARCQVTLRDLFDRAVVLSHYIHN   60
                   . *. * : *                    .     : : :*::**    *. :*.

E. akaara        LSTTLSHDLDFNFPPIGRMAMPRPSMCHTSSLQTPSDKEQALQVSESDLLSLARSLLQAW  103
E. coioides      LSTTLSHDLDFNFPPIGRMAMPRPSMCHTSSLQTPSDKEQALQVSESDLLSLARSLLQAW  103
A. schlegelii    LSTTLTQDLNNHIPPVGWMMMPRPSTCHTSSLQTPNDKEQALQLSESDLLSLARSLLQAW  103
B. melanopus     LSTMLTRELDSHFPPIGRMIMPRPAMCHTSSLQTPNDKEQALQVSESDLLSGKMGLLRAW  103
O. niliticus     LSSTLTQELDSHFPPIGRVIMPRPAMCHTSSLQTPIDKDQALQVSESDLMSLARSLLQAW  103
C. autumnalis    LSTSLTKDLDSHFPPMGRVMMPRPSMCHTSSLQTPKDKEQALRVSENELISLARSLLLAW  102
O. mykiss        LSTSLTKDLDSHFPPMGRVMMPRPSMCHTSSLQTPKDKEQALKVSENELISLARSLLLAW  102
P. hamadryas     LSSEMFSEFDKRYTHGRGFITRAINSCHTSSLPTPEDKEQAQQINQKDFLSLIVSILRSW  119
H. sapiens       LSSEMFSEFDKRYTHGRGFITKAINSCHTSSLATPEDKEQAQQMNQKDFLSLIVSILRSW  120
                 **: :  ::: .        .         ****** ** **:** ::.:.::**  *:* :*

E. akaara        ADPLVVLSTSANTLPHPAQSTISNKIQELQEHTKSLGDGLDILSGKMGPAAQLISSLPYR  163
E. coioides      ADPLVVLSTSANTLPHPAQSTISNKIQELQEHTKSLGDGLDILSGKMGPAAQLISSLPYR  163
A. schlegelii    QDPLVVLSNSANSLVHPSQSSISNKIRELQEHSKSLGDGLDILSGKMGPEAQAISSLPYR  163
B. melanopus     VDPLLVLSSSANTLPHPAQNSISNKIQELQEHSKTLGDGLNILSGKMGPAAQTMSLLPYR  163
O. niliticus     SDPLVVLSSSASTLPHPAQSTIFNKIQEMQQYSKSLKDGLDVLSSKMGSPAQAITSLPYR  163
C. autumnalis    NDPLLLLSSEAPTLPHPSNGDISSKIRELQDYSKSLGDGLDILVNKMGPSSQYISSIPFK  162
O. mykiss        NDPLLLLSSEAPTLPHPSNGDISSKIRELQDYSKSLGDGLDIMVNKMGPSSQYISSIPFK  162
P. hamadryas     NEPLYHLVTEVRGMEEA-PEAILSKAVEIEEQTKRLLEGMELIVSQVHPETKENEIYPVW  178
H. sapiens       NEPLYHLVTEVRGMQEA-PEAILSKAVEIEEQTKRLLEGMELIVSQVHPETKENEIYPVW  179
                 :** * ... : .       * .* *::: :* * :*::::. .::    ::      *

E. akaara        GGND--IGQDRISKLVNFNFLLSCFRRDSHKIDSFLKVLRCRAAKLRPEMC          212
E. coioides      GGSD--IGQDRISKLVNFNFLLSCFRRDSHKIDSFLKVLRCRAAKLRPEMC          212
A. schlegelii    GSND--IGEDSISKLTNFHFLLSCFRRDSHKIDSFLNVLRCRAAKLQPEMC          212
B. melanopus     GGND--LGQDKISKLINFQFLLSCFRRDSHKIDSFLKVLRCRAAKMQPEMC          212
O. niliticus     GGTN--LGHDKITKLINFNFLLSCLRRDSHKIDSFLKVLRCRAAKMQPEMC          212
C. autumnalis    GG-D--LGNDKTSRLINFHFLMSCFRRDSHKIDSFLKVLRCRATKMRPETC          210
O. mykiss        GG-D--LGNDKTSRLINFHFLMSCFRRDSHKIDSFLKVLRCRATKMRPEAC          210
P. hamadryas     TGLPSLQMADEESRLSAYYNLLHCLRRDSHKIDNYLKLLKCRIIHNN--NC          227
H. sapiens       SGLPSLQMADEESRLSAYYNLLHCLRRDSHKIDNYLKLLKCRIIHNN--NC          228
                   .         *  ::* :  *: *:********.:*::*:**  : . *
```

Fig. 2. Multiple alignment of amino acid sequences of prolactin in red-spotted grouper *Epinephelus akaara* compared to those of other teleosts. Prolactin sequences adapted from *Epinephelus coioides* (AAR97730), *Acanthopagrus schlegelii* (AAX21764), *Amphiprion melanopus* (AEB00558), *Oreochromis niloticus* (NP_001266715), *Coregonus autumnalis* (CAA80660), *Oncorhynchus mykiss* (NP_001118205), *Papio hamadryas* (ADG56475), and *Homo sapiens* (AAH88370) were aligned using ClustalW. The putative signal peptide sequence is marked by a clear gray-lined box and pairs of cysteine residues are shaded in gray. Identical amino acids among proteins are indicated by asterisks.

similar in structure to mouse growth hormone (Manzon, 2002).

Expression profile of PRL transcript

RT-PCR was performed to analyze the tissue-specific expression pattern of the PRL gene. The gene was expressed mostly in the brain, including the pituitary gland, of RG, with little if any expression in the other tissues examined (Fig. 4A). A similar pattern of expression was also observed in the hybrid grouper generated between *E. akaara* × *E. lanceolatus* (personal observation). These observations are consistent with those in other species, including tiger pufferfish *Takifugu rubripes*, black porgy *A. schlegelii*, and starry flounder *Platichthys stellatus* (Chang et al., 2007; Lee et al., 2006; Noh et al., 2013a). On the other hand, in the orange-spotted grouper *E. coioides*, goldfish *Carassius auratus*, and olive flounder *P. olivaceus*, PRL gene expression has also been detected in other tissues, including the testis, ovary, liver,
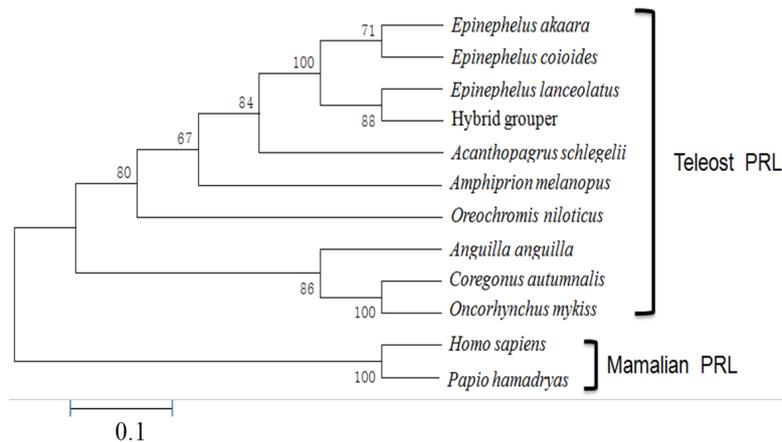
Fig. 3. Construction of neighbor-joining phylogenetic tree based on the amino acid sequences of prolactin. Bootstrap values are indicated for each node. Taxonomic groups are indicated on the right. Prolactin sequences from *Epinephelus coioides* (AAR97730), *Acanthopagrus schlegelii* (AAX21764), *Amphiprion melanopus* (AEB00558), *Oreochromis niloticus* (NP_001266715), *Coregonus autumnalis* (CAA80660), *Epinephelus lanceolatus* (XP_033500110), *Anguilla anguilla* (XP_035254403), and *Oncorhynchus mykiss* (NP_001118205) were used for comparison. Prolactin sequences from *Papio hamadryas* (ADG56475) and *Homo sapiens* (AAH88370) were included as outgroups.
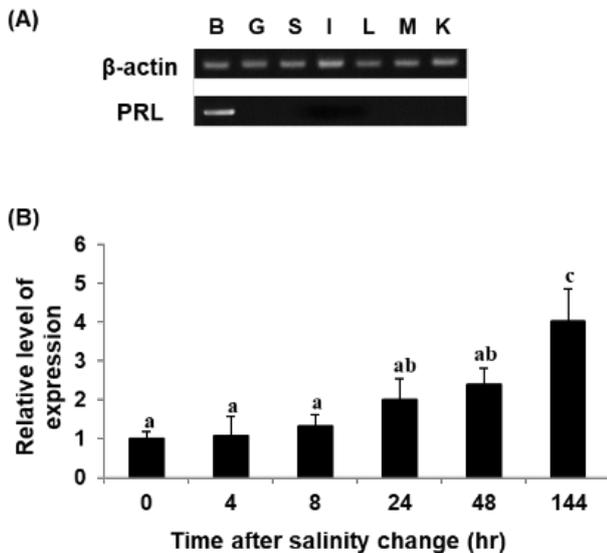
Fig. 4. Tissue-specific distribution of prolactin from red-spotted grouper *Epinephelus akaara* and level of the PRL transcript from brain tissue following a shift to the salinity of 8 psu. (A) RT-PCR was performed with cDNA prepared from RNA isolated from various tissues, including the brain around the pituitary gland (B), gill (G), stomach (S), intestine (I), liver (L), muscle (M), and kidney (K). Detected bands were consistent with the sizes of PRL (~639 bp) expected from the positions of the primers used for amplification. (B) Level of the PRL transcript in brain including the pituitary gland of red-spotted grouper following exposure to low-salinity conditions ($p < 0.05$).

kidney, spleen, gill, and muscle. This may be due to differences in the level of exclusive dissection of the pituitary resulting in relatively less drastic differences in the level of detection under the conditions tested. However, such a difference could also be plausibly explained by different species of fish in different stages of development in different habitats, considering that there may be more than 300 potential roles of PRL in vertebrates (Si et al., 2021).

The PRL transcript was also analyzed upon shifting to 8 psu. The level of PRL transcript increased upon exposure of RG to low salinity under the conditions tested (Fig. 4B), consistent with observations in other fish species. However, the experimental fish adapted to 8 psu for 144 h showed the reascent of PRL gene expression, and this fact differs from the results of other fish species (Noh et al., 2013a). This may have been because not all juvenile *E. akaara* survived under salinity of 8 psu, suggesting continuous stressful conditions and vulnerability to low salinity; this is different from GG, which can survive low-salinity shock (Lim et al., 2016).

Differences in the size of the first intron of grouper PRL genes

Analysis of genomic DNA and cDNA templates indicated a similar structure of the PRL gene consisting of five exons interrupted by four introns in the grouper species examined. The PRL gene in giant grouper *E. lanceolatus* also consists of five
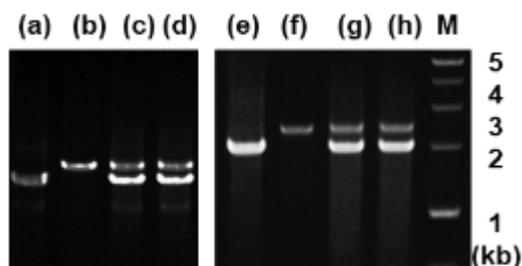
Fig. 5. PCR amplification of the PRL gene from genomic DNA templates isolated from red-spotted grouper (a, e), giant grouper (b, f), and two hybrids generated from red-spotted grouper and giant grouper (c, d, g, f). Primers used in each PCR corresponded to the first and second exons (a–d) and the third and fourth exons (e–h) encompassing the first and third introns, respectively.
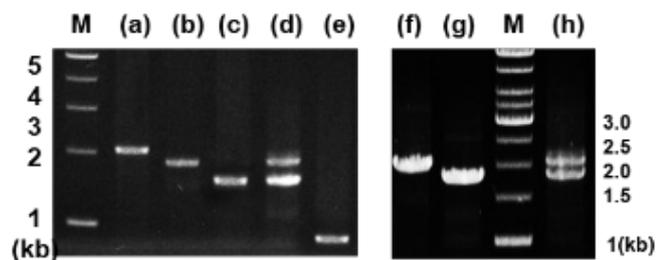
Fig. 6. Analysis of the PCR-amplified fragments encompassing the first intron of the PRL gene from various grouper species. Genomic DNA templates isolated from kelp grouper (a, f), giant grouper (b, g), red-spotted grouper (c), hybrid between giant grouper × red-spotted grouper (d), seven-band grouper (e), and the hybrid between giant grouper × kelp grouper (h). PCR was performed with primers PRL-E1F and PRL-E2R corresponding to the first and second exons, respectively, followed by agarose gel electrophoresis.

exons of 43, 113, 108, 183, and 192 bp leading to an open reading frame of 212 aa residues, similar to RG (personal observations). While the overall sizes of the exons in the PRL genes of the two grouper species were similar to each other and to other fishes, the sizes of the four introns in the PRL gene of the giant grouper (1,671, 90, 2,325, 1,277 bp listed from 5′ to 3′) were different from those of RG. To examine whether the differences in intron sizes can be used to distinguish among the species often indiscriminately identified as valuable fish, the sizes of the corresponding regions of the PRL gene in other groupers were compared following amplification. These were first tested with red-spotted grouper, giant grouper, and the hGR hybrid gen-

erated between them (Lim et al., 2016) using primers designed based on the conserved sequences of exons of RG to amplify the region encompassing each intron. There were clear differences in the sizes of the amplified fragments encompassing the first intron (Fig. 5) among RG, GG, and the hybrid. In particular, the presence of a double band corresponding to the size of the parental copy indicated its applicability for identifying artificially produced hybrids, some of which have phenotypes that are indistinguishable from the parental species. These results indicate that this method may be useful for distinguishing not only different species but also hybrids generated between them. This study was further extended to other *Epinephelus* species, including seven-band grouper, and the hybrids produced between giant grouper, which has the fastest growth rate, and other groupers known to have good meat quality. The distinct differences in the sizes of the amplified products (Fig. 6) indicated that most of the commercially important grouper species could be clearly identified by EPIC of the PRL gene. The method is advantageous compared to 16S rRNA-based detection of three species of grouper, i.e., *E. bruneus*, *E. septemfasciatus*, and *Niphon spinosus* (Park et al., 2013), requiring primers for each fish species. Our method may not only facilitate grouper species identification but also help expand the horizons of a rather sporadically used intron-based fish-identification method.

Competing interests

No potential conflict of interest relevant to this article was reported.

Availability of data and materials

Upon reasonable request, the datasets of this study can be available from the corresponding author.

Ethics approval and consent to participate

This study conformed to the guidance of animal ethical treatment for the care and use of experimental animals.

ORCID

Bok-Ki Choi	https://orcid.org/0009-0006-3427-1150
Gyeong-Eon Noh	https://orcid.org/0000-0001-9561-2015
Yeo-Reum Kim	https://orcid.org/0000-0003-0318-3319
Jun-Hwan Byun	https://orcid.org/0000-0001-8446-9233
Han Kyu Lim	https://orcid.org/0000-0001-8522-9901
Jong-Myoung Kim	https://orcid.org/0000-0003-1005-0952

References

Astola A, Ortiz M, Calduch-Giner JA, Pérez-Sánchez J, Valdivia MM. Isolation of *Sparus auratus* prolactin gene and activity of the cis-acting regulatory elements. Gen Comp Endocrinol. 2003;134:57-61.

Breves JP, Popp EE, Rothenberg EF, Rosenstein CW, Maffett KM, Guertin RR. Osmoregulatory actions of prolactin in the gastrointestinal tract of fishes. General and Comparative Endocrinology. 2020;113589.

Chang YS, Huang FL, Lo TB. Molecular cloning of silver carp and bighead carp prolactin. Gen Comp Endocrinol. 1992;87:260-5.

Chang YJ, Min BH, Choi CY. Black porgy (*Acanthopagrus schlegeli*) prolactin cDNA sequence: mRNA expression and blood physiological responses during freshwater acclimation. Comp Biochem Physiol B Biochem Mol Biol. 2007;147:122-8.

Chao SC, Pan FM, Chang WC. Nucleotide sequence of carp prolactin cDNA. Nucleic Acids Res. 1988;16:9350.

Cooke NE, Coit D, Shine J, Baxter JD, Martial JA. Human prolactin: cDNA structural analysis and evolutionary comparisons. J Biol Chem. 1981;256:4006-16.

Eschmeyer WN, Fricke RL, van der Laan R. Genera in the family Serranidae. Catalog of Fishes. San Francisco, CA: California Academy of Sciences; 2010.

Gong N, Ferreira-Martins D, Norstog JL, McCormick SD, Sheridan MA. Discovery of prolactin-like in lamprey: role in osmoregulation and new insight into the evolution of the growth hormone/prolactin family. Proc Natl Acad Sci USA. 2022;119:e2212196119.

Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proc R Soc Lond B. 2003;270:313-21.

Inoue H, Nojima H, Okayama H. High efficiency transformation of *Escherichia coli* with plasmids. Gene. 1990;96:23-8.

Kang MJ, Noh CH, Kim JH, Park JY, Park DW. The embryonic development and hatchability of two hybrids with giant grouper female: giant grouper ♀×kelp grouper ♂ and giant grouper ♀×red-spotted grouper ♂. Korean J Fish Aquat Sci. 2020;53:557-62.

Lee CH, Hur SW, Kim BH, Soyano K, Lee YD. Induced maturation and fertilized egg production of the red-spotted grouper, *Epinephelus akaara*, using adaptive physiology of photoperiod and water temperature. Aquac Res. 2020;51:2084-90.

Lee KM, Kaneko T, Aida K. Prolactin and prolactin receptor expressions in a marine teleost, pufferfish *Takifugu rubripes*. Gen Comp Endocrinol. 2006;146:318-28.

Li C, Riethoven JJM, Ma L. Exon-primed intron-crossing (EPIC) markers for non-model teleost fishes. Evol Biol. 2010;10:90.

Lim SG, Han SB, Lim HK. Effects of salinity on the growth, survival and stress responses of red spotted grouper *Epinesphelus akaara* and hybrid grouper *E. akaara* ♀ × *E. lanceolatus* ♂. Korean J Fish Aquat Sci. 2016;49:612-9.

Mai W, Liu P, Chen H, Zhou Y. Cloning and immune characterization of the c-type lysozyme gene in red-spotted grouper, *Epinephelus akaara*. Fish Shellfish Immunol. 2014;36:305-14.

Manzon LA. The role of prolactin in fish osmoregulation: a review. Gen Comp Endocrinol. 2002;125:291-310.

Monaghan MT, Balke M, Gregory TR, Vogler AP. DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. Philos Trans R Soc Lond B Biol Sci. 2005;360:1925-33.

Noh CH. Hatchability of fertilized eggs from grouper (subfamily Epinephelinae) hybrids in Korea: a mini review for selection of commercially promising cross combinations. Korean J Fish Aquat Sci. 2020;53:479-85.

Noh GE, Lim HK, Kim JM. Characterization of genes encoding prolactin and prolactin receptors in starry flounder *Platichthys stellatus* and their expression upon acclimation to freshwater. Fish Physiol Biochem. 2013a;39:263-75.

Noh GE, Rho S, Chang YJ, Min BH, Kim JM. Gene encoding prolactin in cinnamon clownfish *Amphiprion melanopus* and its expression upon acclimation to low salinities. Aquat Biosyst. 2013b;9:1.

Park YC, Jung YH, Kim MR, Shin JH, Kim KH, Lee JH, et al. Development of detection method for *Niphon spinosus*, *Epinephelus bruneus*, and *Epinephelus septemfasciatus* using 16S rRNA gene. Korean J Food Sci Technol. 2013;45:1-7.

Quérat B, Cardinaud B, Hardy A, Vidal B, D'Angelo G. Se-

quence and regulation of European eel prolactin mRNA. Mol Cell Endocrinol. 1994;102:151-60.

Rentier-Delrue F, Swennen D, Prunet P, Lion M, Martial JA. Tilapia prolactin: molecular cloning of two cDNAs and expression in *Escherichia coli*. DNA. 1989;8:261-70.

Rimmer MA, Glamuzina B. A review of grouper (family Serranidae: subfamily Epinephelinae) aquaculture from a sustainability science perspective. Rev Aquac. 2019;11:58-87.

Sadovy Y, Liu M, Amorim P, Choat JH, Law C, Ma K, et al. *Epinephelus akaara*. The IUCN red list of threatened species. Gland: International Union for Conservation of Nature (IUCN); 2018.

Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4:406-25.

Si Y, Li H, Gong X, Bao B. Isolation of prolactin gene and its differential expression during metamorphosis involving eye migration of Japanese flounder *Paralichthys olivaceus*. Gene. 2021;780:145522.

Song S, Trinh KY, Hew CL, Hwang SJ, Belkhode S, Idler DR. Molecular cloning and expression of salmon prolactin cDNA. Eur J Biochem. 1988;172:279-85.

Swennen D, Poncelet AC, Sekkali B, Rentier-Delrue F, Martial JA, Belayew A. Structure of the tilapia (*Oreochromis mossambicus*) prolactin I gene. DNA Cell Biol. 1992;11:673-84.

Varma S, Kwok S, Ebner KE. Cloning and nucleotide sequence of ovine prolactin cDNA. Gene. 1989;77:349-59.

Xiong F, Chin RA, Hew CL. A gene encoding chinook salmon (*Oncorhynchus tschawytscha*) prolactin: gene structure and potential cis-acting regulatory elements. Mol Mar Biol Biotechnol. 1992;1:155-64.

Yamaguchi K, Specker JL, King DS, Yokoo Y, Nishioka RS, Hirano T, et al. Complete amino acid sequences of a pair of fish (tilapia) prolactins, $tPRL_{177}$ and $tPRL_{188}$. J Biol Chem. 1988;263:9113-21.

Zhang W, Tian J, Zhang L, Zhang Y, Li X, Lin H. cDNA sequence and spatio-temporal expression of prolactin in the orange-spotted grouper, *Epinephelus coioides*. Gen Comp Endocrinol. 2004;136:134-42.