J. Biochem. 125, 476-486 (1999)

Sequence Analysis of cDNA and Genomic DNA, and mRNA Expression of the Medaka Fish Homolog of Mammalian Guanylyl Cyclase C[1]

Tsuyoshi Mantoku, Ryouji Muramatsu, Mina Nakauchi, Sayaka Yamagami, Takehiro Kusakabe, and Norio Suzuki[2]

Division of Biological Sciences, Graduate School of Science, Hokkaido University, Kita 10, Nishi 8, Kita-ku, Sapporo 060-0810

Received October 19, 1998; accepted November 20, 1998

We isolated the cDNA and genomic DNA encoding a membrane guanylyl cyclase of medaka fish (designated as *OlGC6*), and determined their complete nucleotide sequences. The open reading frame for *OlGC6* cDNA predicted a protein of 1,075 amino acids. Phylogenetic analysis indicated that *OlGC6* is a member of the enterotoxin/guanylin receptor family. We also determined the partial genomic structure of the gene of another membrane guanylyl cyclase of medaka fish, *OlGC2*, which is a member of the natriuretic peptide receptor family. The intron positions relative to the protein-coding sequence are highly conserved in the intracellular domains of *OlGC6*, *OlGC2*, mammalian *GC-A*, and *GC-E*. Despite their divergent primary structures, some intron positions also seem to be conserved in the extracellular domains of different membrane guanylyl cyclase genes. Northern blot analysis demonstrated that an *OlGC6* transcript of 3.9 kb is only present in the intestine, while reverse transcription (RT)-PCR analysis demonstrated that the *OlGC6* transcript is present in the kidney, spleen, liver, pancreas, gallbladder, ovary, testis, brain, and eye. RT-PCR also demonstrated that *OlGC6* is only expressed zygotically and that transcripts are present from 1 day after fertilization, *i.e.* long before the intestinal tissues begin to develop.

Key words: development, exon-intron organization, gene expression, medaka fish, membrane guanylyl cyclase.

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

Guanylyl cyclase (GC) catalyzes the conversion of GTP to cGMP, which is a ubiquitous second messenger in intracellular signaling cascades and is responsible for a wide variety of physiological responses (*1*). There are two forms of guanylyl cyclase, a soluble form (soluble GC) and a membrane form (membrane GC). A soluble GC is a heterodimer composed of α and β subunits (*2*), and is activated by nitric oxide (NO) or carbon monoxide (*1*). Both subunits possess a catalytic domain in the C-terminal part, which is conserved in both membrane GCs and adenylyl cyclases (*3*). A membrane GC consists of a single polypeptide with a molecular mass of about 150-200 kDa. One of the richest sources of membrane GC is sea urchin spermatozoa, in which the enzyme appears to serve as a cell surface receptor for a small egg peptide named resact (SAP-IIA), which possesses both chemokinetic and chemoattractant properties (*4-6*), and causes marked elevation of the cGMP level in spermatozoa. A cDNA encoding a membrane GC was first isolated from a sea urchin testis cDNA library (*7*). Since then, a number of membrane GCs have been isolated from a number of animal tissues. These include a natriuretic peptide receptor and an enterotoxin/guanylin receptor.

In mammals, cDNA clones for seven membrane GCs (GC-A, GC-B, GC-C, GC-D, GC-E, GC-F, and GC-G) have been isolated and characterized (*8-14*). In all cases, the predicted protein contains an extracellular domain, a kinase-like domain, and a cyclase catalytic domain. GC-A and GC-B are activated through the binding of natriuretic peptides, such as ANP, BNP, and CNP, to their extracellular domain (*8, 9, 15-17*). GC-D, GC-E, and GC-F are activated by an intracellular activator (*18*). GC-C has been suggested to be the receptor for *Escherichia coli* heat-stable enterotoxin (STa) (*10, 19, 20*). Two endogenous peptides, guanylin and uroguanylin, have been identified and suggested to be normal ligands for GC-C (*21, 22*). Both peptides stimulate the cyclase and compete in STa binding to the enzyme, although there is no direct evidence to indicate that GC-C is the principal or only protein that binds to these peptides. The cyclase is expressed in the mammalian intestine at a high level, as well as in other tissues at lower levels (*10, 23-27*). Since all mammals appear to express

[1] This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan (Nos. 08454267 and 10102001). The nucleotide sequence data reported in this paper have been submitted to the DDBJ, EMBL, and GenBank nucleotide sequence databases, and are available under accession numbers AB007192 (cDNA of *OlGC6*), AB016081 (genomic DNA of *OlGC6*), and AB016082 (genomic DNA of *OlGC2*).
[2] To whom correspondence should be addressed. Tel: +81-11-706-4908, Fax: +81-11-746-1512, E-mail: norio-s@sci.hokudai.ac.jp
Abbreviations: GC, guanylyl cyclase; cGMP, guanosine 3',5'-cyclic monophosphate; *OlGC*, gene encoding *Oryzias latipes* membrane guanylyl cyclase; ORF, open reading frame; RT-PCR, reverse transcription-polymerase chain reaction; UTR, untranslated region.

© 1999 by The Japanese Biochemical Society.

GC-C and are susceptible to STa-induced acute secretory diarrhea, it is assumed that the evolutionary pressure to retain a functional receptor/GC-C is due to its having an essential role in the normal regulation of intestinal or other tissues, and/or its requirement as an essential protective agent against, for example, certain enteric infections. The gene encoding GC-C, *gucy2*, has been mapped to distal chromosome 6 in the mouse, and chromosome 12p12 in man (*28*), for which no genetic diseases have yet been described. GC-C deficient mice are viable, fertile, and develop normally, but are resistant to STa-induced diarrhea and to enterotoxigenic bacteria that produce STa (*29, 30*).

Vertebrate membrane GCs can be grouped into three major subfamilies, (i) natriuretic peptide receptors, (ii) enterotoxin/guanylin receptors, and (iii) sensory organ-specific GCs (*31*). While the exon-intron organizations of a natriuretic peptide receptor GC gene (*32, 33*) and sensory organ-specific GC genes (*34, 35*) have been reported, the entire exon-intron organization of the enterotoxin/guanylin receptor/GC-C gene has not yet been clarified.

In *Drosophila* and medaka fish, some membrane GC transcripts as well as a soluble GC are present in unfertilized eggs, and in embryos of early developmental stages (*31, 36-38*). These facts suggest that the cGMP signaling pathway mediated by membrane and soluble GCs is involved in the early development of these animals. Despite the rapidly accumulating information concerning the tissue/cell distributions and the physiological roles of membrane GCs, there have been few reports regarding the importance of the cGMP pathway during vertebrate development. In this study, in order to determine the developmental roles of membrane GCs in vertebrates, we have characterized a homolog of mammalian GC-C in the medaka fish, *Oryzias latipes*, a species allowing both classical and molecular genetic analyses. We report here the primary structures of cDNA and genomic DNA encoding a fish homolog of mammalian GC-C, and the expression of this homolog in adult tissues and during embryogenesis. We also discuss the evolution of membrane GC genes based on the exon-intron organizations of various membrane GCs in vertebrates.

MATERIALS AND METHODS

Animals and Embryos—Mature adults of the orange-red variety of medaka fish, *O. latipes*, were purchased from a local dealer. They were kept in indoor tanks under artificial reproductive conditions (10 h dark, 14 h light, 27°C) and fed on TetraMin flakes (TetraWerke). Naturally spawned and fertilized eggs were collected, and the embryos were cultured as described previously (*31*). The developmental stage was expressed in days and the day of fertilization was referred to as day 0. Hatching usually occurred on day 10.

Preparation of RNA, cDNA Library Construction, and Screening—Total RNA was extracted according to the AGPC method (*39*). Poly(A)-rich RNA was isolated using Oligotex-dT30<Super> (Roche), according to the manufacturer's protocol. Using poly(A)-rich RNA isolated from the intestines of mature adult medaka fish, an oligo-dT primed cDNA library was constructed in the λZAP II vector using a ZAP-cDNA Synthesis Kit (Stratagene).

Total RNA (3 μg) from rat intestine was used as the template to synthesize the first strand of cDNA using an oligo(dT) primer according to the manufacturer's protocol (SuperScript Preamplification System for First Strand cDNA Synthesis; Life Technologies). A 1,357-bp fragment of rat GC-C cDNA (nucleotides 1352-2729) was amplified by PCR (30 cycles; 94°C for 30 s, 55°C for 30 s, and 72°C for 3 min) with a pair of oligonucleotide primers, 5'-ATTGCC-CTCCTTGTGCTCAGA-3' and 5'-CAAAGGTCCCCATG-AAGCTGA-3', cloned into a plasmid vector, pBluescript II KS(−) (Stratagene), and then used as a probe for the screening of the cDNA library. Approximately 4.0×10^5 plaques were screened with the ^{32}P-labeled random-primed cDNA probe as described previously (*31*).

5'-Rapid Amplification of cDNA Ends (5'-RACE)—To determine the full-length cDNA sequence of *OlGC6*, a 5'-portion of the cDNA was amplified by the 5'-RACE method (*40*) using the 5'-RACE System for Rapid Amplification of cDNA Ends, Version 2.0 (Life Technologies). One hundred nanograms of poly(A)-rich RNA from the intestine was reverse-transcribed with a gene-specific antisense oligonucleotide primer (GSP1, GSP4, or GSP7). The cDNA was tailed with dCTP using terminal deoxynucleotidyl transferase, and then amplified by PCR with an Abridged Anchor Primer (Life Technologies) and another gene-specific antisense oligonucleotide primer (GSP2, GSP5, or GSP8). A one-tenth volume of the PCR product was amplified by PCR with the Abridged Universal Amplification Primer (Life Technologies) and another gene-specific antisense oligonucleotide primer (GSP3, GSP6, or GSP9). Four different PCR conditions were used: for GSP2, GSP8, and GSP9, denaturation at 94°C for 3 min followed by 35 amplification cycles (94°C for 30 s and 68°C for 2 min) and final extension at 72°C for 9 min; for GSP3, 30 amplification cycles (94°C for 30 s, 60°C for 30 s, and 72°C for 1 min) and final extension at 72°C for 9 min; for GSP5, 35 amplification cycles (94°C for 30 s, 61°C for 30 s, and 72°C for 2 min) and final extension at 72°C for 9 min; and for GSP6, 30 amplification cycles (94°C for 30 s, 62°C for 30 s, and 72°C for 2 min) and final extension at 72°C for 9 min. The PCR products were cloned into pBluescript II KS(−) and sequenced. The gene specific primers used were complementary to: nucleotide positions 1843-1861 (GSP1), 1779-1803 (GSP2), 1758-1781 (GSP3), 1175-1192 (GSP4), 1140-1165 (GSP5), 1121-1146 (GSP6), 616-635 (GSP7), 602-629 (GSP8), and 556-583 (GSP9). Each 5'-RACE product overlapped, by 54-158 bp, the 5' end of the clone that had been isolated, and the sequences of these overlapping regions were identical (Fig. 1).

Sequence Comparison—The amino acid sequence of OlGC6 was compared with those of known membrane GCs using DNASIS software (Hitachi Software Engineering). The GenBank/EMBL/DDBJ accession numbers for the sequences used are: X14773, rat GC-A (*8*); M26896, rat GC-B (*9*); M55636, rat GC-C (*10*); L37203, rat GC-D (*11*); L36029, rat GC-E (*12*); L36030, rat GC-F (*12*); M73489, human GC-C (*19*); D17513, pig GC-C (*41*); Z74734, guinea pig GC-C (Kruhoeffer, M., Cetin, Y., Kaempf, U., and Forssmann, W.G., unpublished data); D49837, frog (*Xenopus laevis*) GC-C (MacFarland, R.T., unpublished data); AB000899, OlGC3 (*31*); AB000900, OlGC4 (*31*); AB000901, OlGC5 (*31*); and D21101, sea urchin (*Hemicentrotus pulcherrimus*) spermatozoa GC (*42*).

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

Reverse Transcription-Polymerase Chain Reaction (RT-PCR)—Total RNA (1 μg) was used as the template to synthesize the first strand cDNA using an oligo(dT) primer according to the manufacturer's protocol (SuperScript Preamplification System for First Strand cDNA Synthesis; Life Technologies). The cDNA fragment containing the 3′ untranslated region (UTR) of *OlGC6* was amplified by PCR using a pair of oligonucleotide primers, 5′-ACAACGGAG-AACGT-CCAGCGT-3′ and 5′-GTAACGAGGCAACATG-GCTGC-3′. As an internal control, the cDNA fragment of 3′ UTR of medaka cytoplasmic actin gene *OlCA1* was amplified as described previously (*31*). The PCR conditions used were: denaturation at 95°C for 5 min followed by 25 amplification cycles (95°C for 30 s, 60°C for 30 s, and 72°C for 1 min) and final extension at 72°C for 9 min, using either a 1/10 volume of the cDNA reaction mixture or 1 pg of plasmid DNA (positive control) as a template. To eliminate the possibility of PCR amplification due to contaminating genomic DNA, the same reaction was carried out without reverse-transcriptase as a control experiment. One-tenth (*OlGC6*) or one-twelfth of the PCR products was separated on a 1.2% NuSieve/0.4% SeaKem GTG agarose (FMC) gel. The DNA fragment was transferred to a nylon membrane (Hybond-N+, Amersham) with 0.4 M NaOH as the transferring solution, and the membrane was washed with 2× SSC. The membranes were prehybridized for 2 h at 42°C in 5×SSPE (1×SSPE: 0.15 M NaCl, 8.65 mM sodium phosphate, and 1.25 mM EDTA, pH 7.4) containing 0.5% SDS, 5×Denhardt's solution, 100 μg/ml denatured herring sperm DNA, and 50% formamide. ^{32}P-labeled probes were synthesized by the random primer method. The probe (10^6 cpm/ml) was added to the prehybridization solution, followed by incubation overnight at 42°C. The membranes were washed with 2×SSC containing 0.1% SDS for 40 min at 65°C and then twice with 0.1×SSC containing 0.1% SDS for 40 min at 65°C. The size of the DNA was estimated using 1 kbp ladder (Life Technologies) bands as standards. Imaging of radioactive signals was performed with a FUJIX Bio-Imaging Analyzer BAS2000 (Fuji Photo Film).

Northern Blot Hybridization—Poly(A)-rich RNA (2 μg) or total RNA (30 μg) was separated on a 1% agarose gel containing 6.7% formaldehyde. The RNA was transferred to a nylon membrane (Hybond-N+, Amersham) and probed with random-primed ^{32}P-labeled cDNA (nucleotide positions 3214-3486) of *OlGC6* as described previously (*31*). Imaging of radioactive signals was performed with the FUJIX Bio-Imaging Analyzer BAS2000.

Cloning and Characterization of Genomic DNA Clones—Three degenerate oligonucleotide primers (P2, P6, and P7) were synthesized based on the amino acid sequences of three conserved regions (DIVGFT, DTVNTA, and MPRY-CL; residues 797-802, 913-918, and 905-910, respectively, in OlGC6) in a known membrane GC (*31*). Another oligonucleotide primer (Exon18: 5′-GTCGTTACGCTGCT-CAA-3′) was synthesized based on the nucleotide sequence of exon 18 of *OlGC1*, another membrane GC gene of medaka fish (Takeda, K. and Suzuki, N., unpublished data). These four primers were used to amplify genomic DNA fragments encoding GC by PCR. A 586-bp genomic DNA fragment encoding a different membrane GC of medaka fish, *OlGC2*, was amplified by nested PCR using oligonucleotide primers P2/P6 and then P7/Exon18. Genomic DNA clones containing the *OlGC2* gene were isolated by screening a medaka fish (white strain) genomic DNA library constructed in λFIX II (Stratagene) using the ^{32}P-labeled 586-bp fragment as a probe. Approximately 9×10^5 plaques were screened as described previously (*31*). The isolated genomic DNA clones contained part of the *OlGC2* gene corresponding to exons 18-22 of the rat *GC-A* gene (*32*). An *OlGC2* genomic fragment corresponding to exons 16 and 17 of the rat *GC-A* gene was further obtained by PCR using a pair of oligonucleotide primers (5′-TGGAGG-AGCGGACTCAAG-3′ and 5′-TCCATGGGCGTGCTCT-CT-3′) synthesized based on the nucleotide sequence of *OlGC2* cDNA (Muramatsu, R. and Suzuki, N., unpublished data).

To isolate genomic DNA clones containing the *OlGC6* gene, the medaka fish genomic DNA library was screened with a 197-bp fragment of the *OlGC2*-coding region and an *OlGC6* cDNA fragment (nucleotides 3051-3622) as probes. The *OlGC2* probe encoded the N-terminal part (amino acid residues 1-67 in Fig. 6) of the cyclase catalytic domain, which is highly conserved among known GCs. The probe was labeled with ^{32}P-dCTP by the random primer method and used to screen 4.5×10^5 plaques as described previously (*31*). The *OlGC6* probe was labeled with digoxigenin (DIG)-dUTP using DIG-High Prime (Boehringer Mannheim) and then used to screen approximately 4.5×10^5 plaques, according to the manufacturer's protocol. The phage DNAs of positive clones were purified using a QIAGEN Lambda Midi Kit (Qiagen) and then digested with *Not*I. The insert DNAs were digested with *Bam*HI, *Eco*RI, and/or *Sal*I, and then analyzed by Southern hybridization using *OlGC6* probes. The digested fragments were subcloned into pBluescript vectors (Stratagene). The subcloned genomic DNA fragments were used directly as templates to determine the nucleotide sequences.

Other Methods—The sequences of the cDNA and genomic DNA fragments in pBluescript vectors (Stratagene) were determined by the dideoxy chain termination method

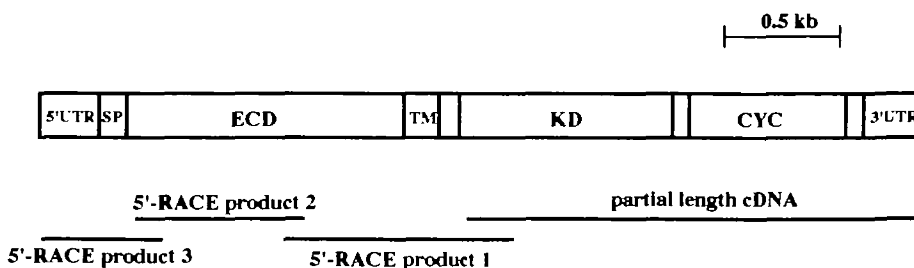

Fig. 1. **Schematic diagram of cDNA clones for *OlGC6*.** The structure predicted for the medaka fish *OlGC6* cDNA is presented at the top. SP, signal peptide; ECD, extracellular domain; TM, transmembrane domain; KD, protein kinase-like domain; CYC, cyclase catalytic domain. The lower lines represent partial length cDNA clones obtained on library screening and with the 5′-RACE method (see "MATERIALS AND METHODS").

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

(*43*) with Applied Biosystems 373A and 377 DNA sequencers or a LI-COR sequencer model 4000L (LI-COR), and analyzed with DNASIS software, GENETYX-MAX/version 9.01 or ATSQ 3.01 (Software Development).

RESULTS

Isolation and Characterization of a cDNA Clone Encoding a Medaka Fish Membrane GC—By repeated screening of a medaka fish intestine cDNA library with a cDNA fragment of rat GC-C as a probe, one positive clone of 2.0

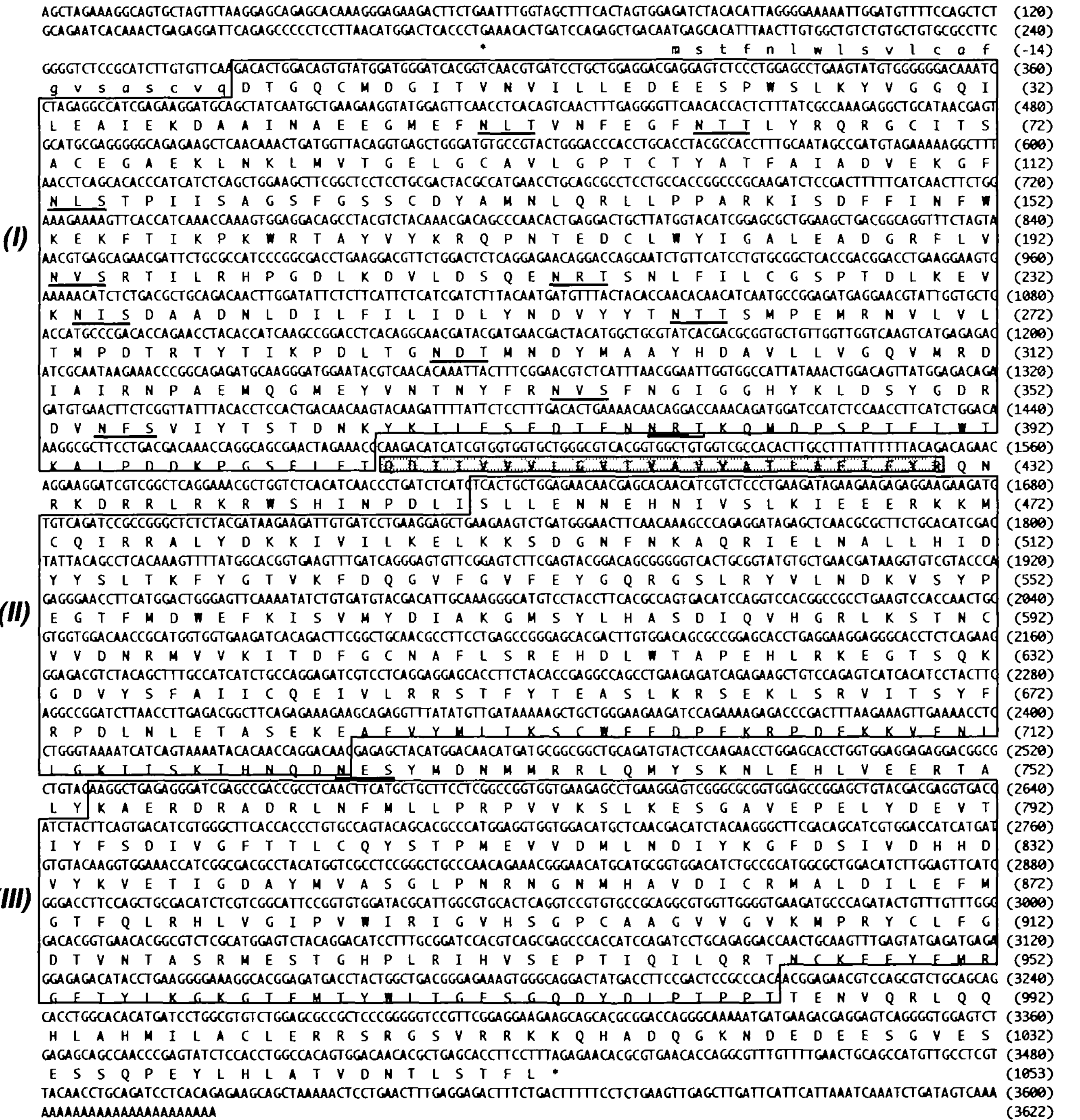

Fig. 2. **The nucleotide and deduced amino acid sequences of *OlGC6* cDNA.** The deduced amino acid sequence is indicated by a single letter code. The signal peptide sequence is indicated by lowercase letters. The amino acids are numbered relative to the predicted signal cleavage site (+1). The predicted transmembrane sequence is indicated by a shaded box. The potential *N*-linked carbohydrate binding sites are underlined. The asterisk (*) denotes a stop codon. Open boxes (I), (II), and (III) indicate the extracellular, kinase-like, and cyclase catalytic domains, respectively.

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

kbp was isolated and designated as *OlGC6*. Comparison of the deduced amino acid sequence of the *OlGC6* cDNA clone with that of enterotoxin/guanylin receptor/GC-C indicated that the cDNA clone lacked the 5′-portion of the coding

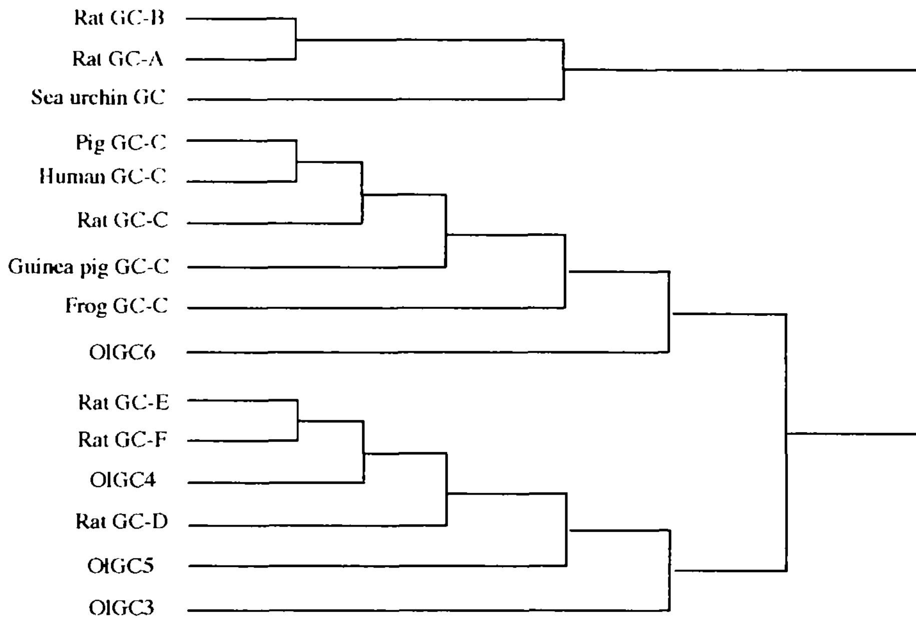

Fig. 3. **Phylogenetic tree and alignment of the amino acid sequences of membrane GCs.** (A) The amino acid sequences of the catalytic domains of various membrane GCs were subjected to phylogenetic analysis. (B) The amino acid sequence of the extracellular domain (412 residues) of OlGC6 was compared with those of various vertebrate GC-Cs. Six conserved cysteine residues are indicated by open boxes. The sources, references, and accession numbers for the GC sequences are given in the text.

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

region. Therefore, the 5′-RACE method (*40*) was adopted to obtain the 5′-portion of *OlGC6* cDNA (Fig. 1). The primary and secondary 5′-RACE reactions produced a 741-bp cDNA fragment and a 617-bp cDNA fragment, respectively. However, neither fragment seemed to contain the initiation codon. The tertiary 5′-RACE reaction resulted in the isolation of a 583-bp cDNA fragment containing the putative initiation codon. The full-length cDNA sequence of *OlGC6* is shown in Fig. 2. The *OlGC6* cDNA consists of a 198-bp 5′ UTR, a 3,228-bp open reading frame (ORF), and a 174-bp 3′ UTR with a poly(A) tail. Termination codons are present in all three frames upstream of the putative initiation codon. In addition, nucleotides around the initiation codon fit the preferred sequence context for the initiation of protein synthesis in eukaryotic mRNAs (*44*). The open reading frame (ORF) of *OlGC6* cDNA encodes a polypeptide of 1,075 amino acids. Hydropathy analysis (data not shown) involving the method of Kyte and Doolittle (*45*) predicted that the OlGC6 polypeptide contains a 22-residue amino-terminal signal sequence (*46*), and that the cleavage of the signal sequence would result in a mature protein of 1,053 amino acids (Fig. 2). Sequence comparison of OlGC6 with known membrane GCs suggested that OlGC6 is closely related to mammalian enterotoxin/guanylin receptors such as rat GC-C, which is specifically expressed in the intestine. The mature OlGC6 protein comprises an extracellular domain (residues 23-429), a single transmembrane domain (residues 430-453), and intracellular protein kinase-like (residues 474-754) and cyclase catalytic (residues 778-1,006) domains, which are highly conserved among invertebrates and vertebrates (Figs. 1, 2, and 3A). However, as shown in Fig. 3B, the amino acid sequence of the extracellular domain of OlGC6 is less conserved except for the relative positions of cysteine, which are more highly conserved among the extracellular domains of enterotoxin/guanylin receptor/GC-Cs than among their protein kinase-like and cyclase catalytic domains.

Expression of the OlGC6 Gene—On Northern blot analysis, an *OlGC6* transcript of 3.9 kb was detected with RNA from the adult intestine, whereas no signal was detected for the eye, brain, or liver (Fig. 4A). As shown in Fig. 4B, however, the *OlGC6* transcript was also detected in the kidney, spleen, liver, pancreas, gallbladder, ovary, testis, brain, and eye of adult fish on RT-PCR analysis, in which the 3′ UTR of the *OlGC6* transcript was amplified and detected by Southern hybridization using a ^{32}P-labeled probe. The signal due to *OlGC6* in the intestine was as strong as that in the pancreas and gallbladder. The signals in tissues other than the intestine, pancreas, and gallbladder were relatively weak and were not detectable on ethidium bromide staining (data not shown). The temporal expression patterns of *OlGC6* during the embryogenesis of medaka fish were examined by RT-PCR analysis. The *OlGC6* transcript was first detected on Day 1, the signal becoming more distinct during later embryogenesis (Fig. 4C).

Genomic Structures of Medaka Fish Membrane GC Genes—Two overlapping clones containing the *OlGC6* gene were isolated from a medaka genomic library. The complete nucleotide sequence of the 18,696-bp genomic region, including the entire *OlGC6* gene, was determined. The sequence analysis showed that the *OlGC6* gene is about 16 kbp in length and contains 27 exons (Fig. 5). The nucleotide sequences of all the splice junctions showed good agreement with the GT/AG rule. It has been demonstrated that all membrane GCs are composed of four domains: an extracellular domain, a transmembrane domain, a kinase-like domain, and a catalytic domain. In the *OlGC6* gene, exons 1-10 correspond to the extracellular domain, exon 11 to the transmembrane domain, exons 12-20 to the kinase-like domain, and exons 21-27 to the catalytic domain (Fig. 5 and Table I). To determine the structural and evolutionary relationships between the *OlGC6* and other membrane GC genes, we isolated genomic DNA clones encoding another membrane GC of medaka fish, OlGC2. The *OlGC2* genomic clones contained the 3′-part of the gene encoding the catalytic domain, and consisted of six exons and five introns (Fig. 5 and Table I). None of the *OlGC2* genomic and cDNA clones obtained so far contains a more 5′-part of the coding region. Comparison of the deduced amino acid sequence of OlGC2 with those of known membrane GCs indicated that OlGC2 is a member of the natriuretic peptide receptor GC subfamily.

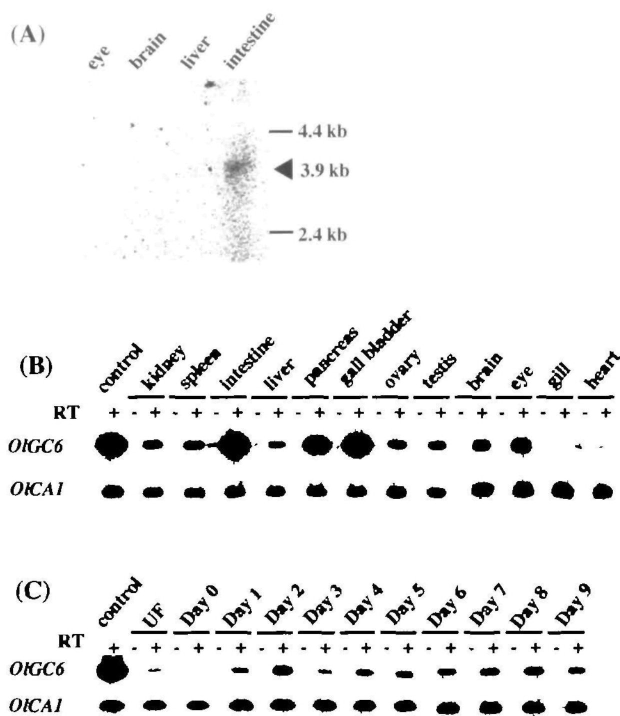

Fig. 4. **Northern blot and RT-PCR analyses of the *OlGC6* transcript in adult medaka fish and during embryogenesis.** (A) Northern blot analysis of total RNA from the brain, eye, liver, and intestine. The radioactive band is indicated by an arrowhead. The positions and sizes of RNA markers are shown on the right. The blot containing 30 μg of total RNA was hybridized with a ^{32}P-labeled cDNA probe for *OlGC6*. Northern blot analysis of poly(A)-rich RNA gave the same results (data not shown). (B) Detection of the *OlGC6* transcript on RT-PCR in various adult organs. (C) Detection of the *OlGC6* transcript in embryos at various developmental stages from UF (unfertilized eggs) to Day 9 (embryos cultured for 9 days after fertilization; hatching stage). A transcript of medaka cytoplasmic actin gene *OlCA1* was amplified as an internal control. RT+ and RT− represent amplification with and without reverse transcriptase, respectively. The PCR products were hybridized with a ^{32}P-labeled cDNA probe for *OlGC6* or *OlCA1*.

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

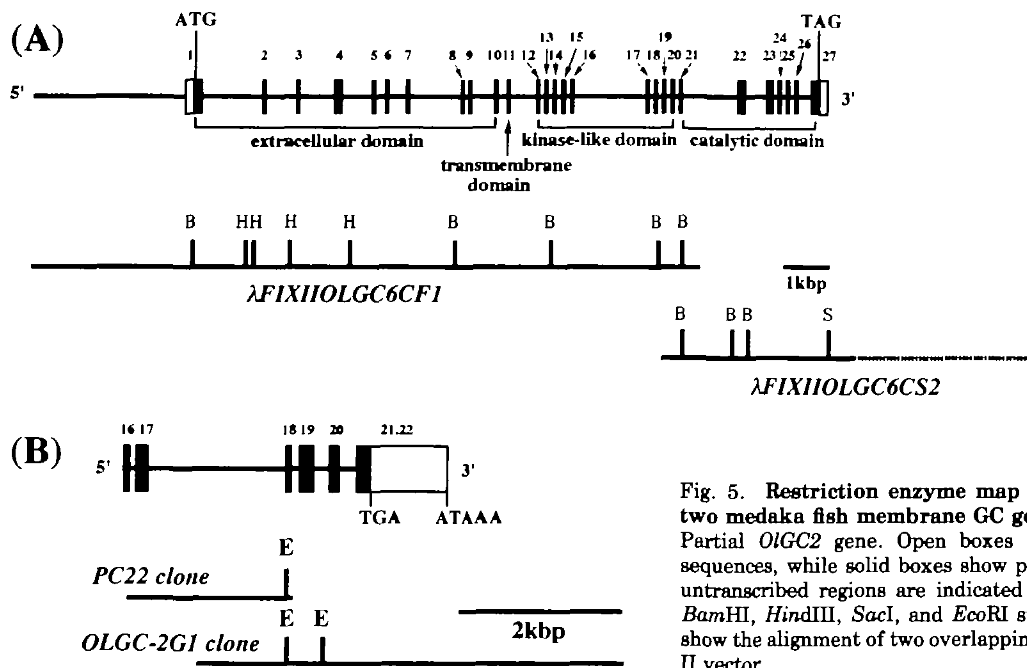

Fig. 5. **Restriction enzyme map and structural organization of two medaka fish membrane GC genes.** (A) Entire *OlGC6* gene. (B) Partial *OlGC2* gene. Open boxes indicate 5′ leader and 3′ trailer sequences, while solid boxes show protein coding regions. Introns and untranscribed regions are indicated by lines. B, H, S, and E denote *Bam*HI, *Hin*dIII, *Sac*I, and *Eco*RI sites, respectively. The lower lines show the alignment of two overlapping genomic DNA clones in the λ FIX II vector.

TABLE I. **Exon-intron organizations of the *OlGC6* and *OlGC2* genes.**

OlGC6			*OlGC2*			Functional domain[b]
Exon/intron No.	Exon size (bp)	Intron size (bp)	Exon/intron No.[a]	Exon size (bp)	Intron size (bp)	
1	403	1,422	–	–	–	ECD
2	113	699	–	–	–	ECD
3	68	830	–	–	–	ECD
4	216	714	–	–	–	ECD
5	119	225	–	–	–	ECD
6	103	452	–	–	–	ECD
7	115	1,154	–	–	–	ECD
8	142	129	–	–	–	ECD
9	83	503	–	–	–	ECD
10	109	255	–	–	–	ECD
11	82	652	–	–	–	TM
12	103	72	–	–	–	KD
13	57	81	–	–	–	KD
14	72	97	–	–	–	KD
15	105	91	–	–	–	KD
16	87	1,668	–	–	–	KD
17	133	65	–	–	–	KD
18	138	68	–	–	–	KD
19	77	80	–	–	–	KD
20	92	80	–	–	–	KD
21	159	1,273	16	>73	70	CYC
22	194	471	17	124	1,722	CYC
23	0	0	18	69	111	CYC
23	175	71	19	175	187	CYC
24	99	79	20	99	222	CYC
25	95	73	21/22	1,069		CYC
26	77	345				CYC
27	382					

[a]Exon numbers of corresponding regions of the mammalian GC-A gene (*32*). [b]Exons coding for the extracellular (ECD), transmembrane (TM), kinase-like (KD), and cyclase catalytic (CYC) domains.

The positions of introns in the *OlGC6* and *OlGC2* genes relative to the primary structure of the protein were compared with those in two mammalian membrane GC genes, *GC-A* (*32*) and *GC-E* (*34*). The intron positions are highly conserved in the cyclase catalytic domain and also conserved to some extent in the kinase-like domain (Fig. 6). On the other hand, the numbers and positions of introns are variable and less-conserved in the genomic DNA region encoding the extracellular domain. However, a few intron positions seem to be conserved between *OlGC6* and mammalian *GC-A* and/or *GC-E* (Fig. 6). The transmembrane domain is encoded by a single exon in each gene.

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

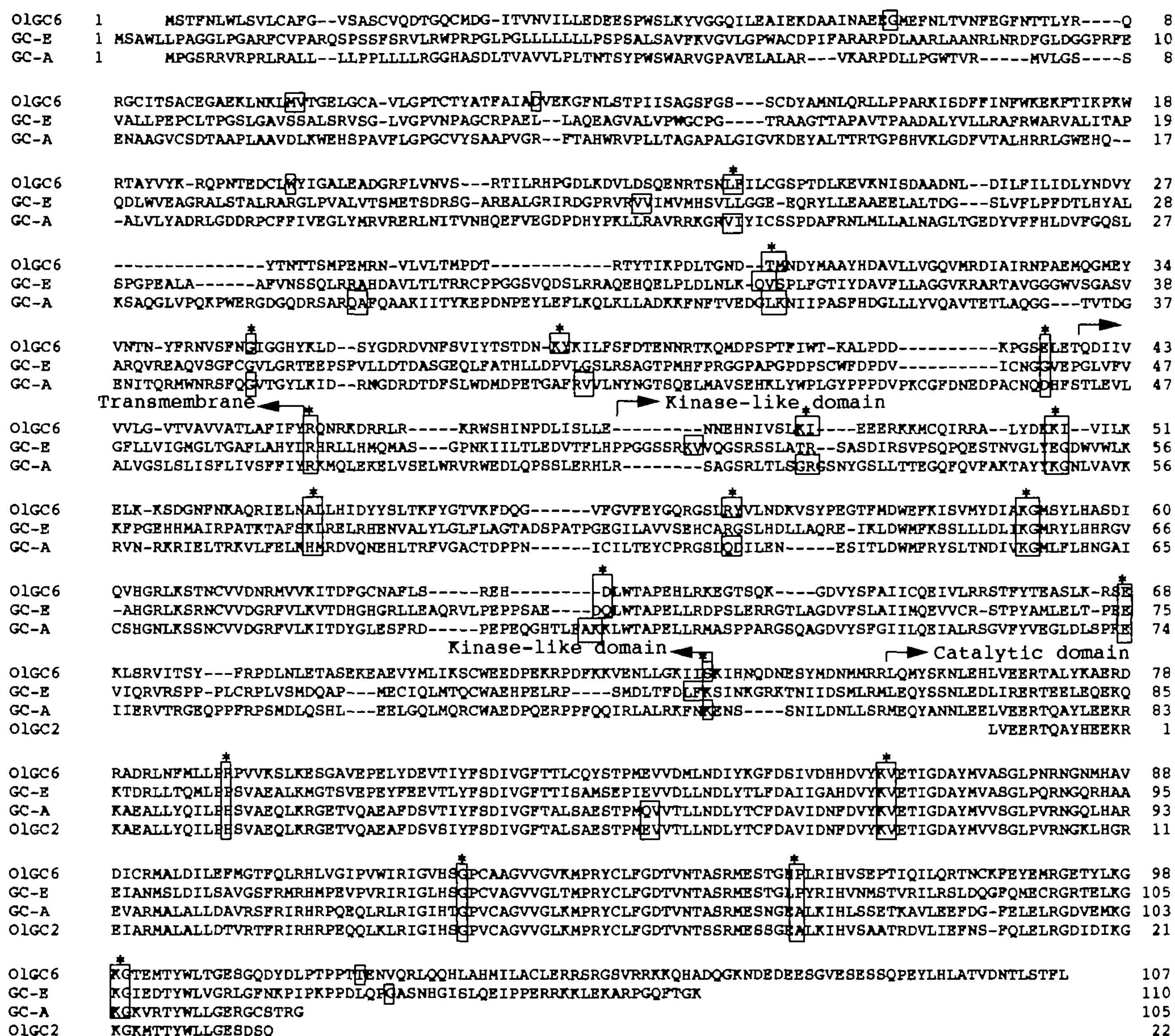

Fig. 6. **Comparison of the relative positions of introns in the *OlGC6*, *OlGC2*, and mammalian GC genes.** The amino acid sequences of OlGC6, OlGC2, rat GC-A, and mouse GC-E were aligned using the ClustalW program (*47*) and MacVector version 6.0.1 (Oxford Molecular). The positions of introns in the genes are indicated by boxed amino acid residues where introns interrupt the coding sequence. Asterisks (*) indicate intron positions conserved in the *OlGC6* and mammalian GC genes.

We determined the nucleotide sequence of the 5′-flanking region of the *OlGC6* gene. Five E-box (CANNTG) motifs, the consensus sequence of binding sites for bHLH transcription factors (*48*), are present within 150 bp upstream of the 5′ end of *OlGC6* cDNA. The 5′ flanking region of the *OlGC6* gene also contains binding consensus sequences for transcription factors known to be expressed in the mammalian intestine, such as Cdx-2, a member of the caudal family of homeodomain proteins (*49*) and of the GATA family of transcription factors (*50*).

DISCUSSION

In vertebrates, three major groups of membrane GCs have been demonstrated: (i) natriuretic peptide receptors, (ii) enterotoxin/guanylin receptors, and (iii) sensory organ-specific GCs (*31*). In the present study, we identified OlGC6, a homolog of enterotoxin/guanylin receptor/GC-C, in the medaka fish intestine. Northern blot analysis indicated that *OlGC6* is expressed predominantly in the adult intestine. On the other hand, RT-PCR analysis demonstrated that *OlGC6* is expressed abundantly not only in the intestine but also in many other organs, such as the kidney, spleen, liver, pancreas, gallbladder, ovary, testis, brain, and eye. This suggests that a member of the heat-stable enterotoxin/guanylin receptor/GC-C subfamily plays previously unrecognized roles in these organs. STa binding has been detected in many extraintestinal organs of the North American opossum, including the liver and kidney, although GC-C has not been directly demonstrated as a binding protein (*25*). In the rat, *GC-C* mRNA has been reported to be expressed in the testis and placenta, and in

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

the fetal, neonatal, and regenerating liver (26, 27). Using RT-PCR, however, Carrithers *et al.* (51) demonstrated that human *GC-C* mRNA was only expressed in the intestinal tract and in tumors derived from intestinal tissues. On the other hand, Gao *et al.* (52) found mRNA for GC-C was expressed in rat pancreas and highly expressed in cultured insulinoma cells. Future studies on medaka fish will elucidate the role of enterotoxin/guanylin receptor/GC-C in extraintestinal tissues.

There have been few reports on the expression patterns of heat-stable enterotoxin/guanylin receptor/GC-C genes during vertebrate development. Our results demonstrated that the transcript of *OlGC6* is present in embryos on Day 1. The intestinal tissues develop at much later stages, suggesting that *OlGC6* plays a role in early embryos different from that in the adult intestine. The cGMP signaling pathway mediated by OlGC6 may be involved in the early development of fish. In *Drosophila*, membrane GCs are expressed during oogenesis and early development (36, 37). In medaka fish, *OlGC3* and *OlGC5* are expressed both maternally and zygotically, while *OlGC4* is only expressed zygotically (31). In mice, expression of *GC-C* is seen by embryonic day 12 in the developing intestine and is also detected transiently in the neonatal liver (53). The cGMP signaling pathways mediated by membrane GCs that regulate oogenesis and early development may thus have been conserved in vertebrates and invertebrates.

In the present study, we characterized the entire genomic structure of *OlGC6*, including 27 putative exons, 26 introns, and the 5′-flanking region. We also determined the genomic DNA sequence encoding the catalytic domain of OlGC2. The exon-intron organizations of the *OlGC6* and *OlGC2* genes were compared with those of the mammalian *GC-A* and *GC-E* genes. Each of the four functional and structural domains of OlGC6 is encoded by a group of exons. That is, introns are located at the boundaries between different domains, as reported for mammalian GC genes (32-35). This conserved feature suggests that "exon-shuffling" events played an important role in the establishment of the basic structure of the ancestral membrane GC gene.

The intron positions are highly conserved in the genomic region encoding the intracellular domain. In the catalytic domain-coding region, the intron positions are identical in the *OlGC6* and *GC-E* genes, and are also highly conserved in *OlGC6* and *GC-A*. *GC-A* contains one additional intron (intron 17) in the catalytic domain-coding region. An intron at this position is also present in *OlGC2*, consistent with the high degree of sequence similarity between *OlGC2* and mammalian *GC-A*. On the other hand, *OlGC2* lacks an intron close to the 3′ end of the catalytic domain-coding region (intron 25 of *OlGC6* and intron 21 of *GC-A*). Since this intron is conserved among *OlGC6, GC-A*, and *GC-E*, it has been probably lost in the ancestral gene of *OlGC2*. *OlGC6* and mammalian *GC-E* have an intron near the 3′ end of the coding region. The intron positions are also well conserved in the kinase-like domain-coding region. This conservation of the exon-intron organization in the intracellular-coding region suggests a common origin of these domains.

The amino acid sequences of the extracellular domains of membrane GCs are highly divergent and only limited sequence similarity has been found between different subfamilies of membrane GCs. The present study showed that the number of introns is variable in the extracellular domain-coding region of membrane GC genes. For example, seven introns interrupt the extracellular domain of OlGC6, while only two introns are present in the corresponding region of mouse GC-E (34). However, our analysis also revealed that the relative positions of some introns seem to be conserved even in the extracellular domain-coding region among the three major subfamilies of membrane GCs. This implies that the extracellular domains of the three different groups of membrane GCs evolved from a common ancestor. Cysteine residues are conserved in the extracellular domains of nematode and vertebrate GCs (54). The conserved intron positions and cysteine residues in the extracellular domain suggest that all membrane GCs originated from a common ancestral protein consisting of extracellular, transmembrane, protein kinase-like, and catalytic domains.

The authors would like to thank the staff members of the Research Center for Molecular Genetics and the Center for Animal Experiments, Hokkaido University, for allowing the use of equipment for probe radiolabeling and culturing of medaka fishes.

REFERENCES

1. Drewett, J.G. and Garbers, D.L. (1994) The family of guanylyl cyclase receptors and their ligands. *Endocrine Rev.* **15**, 135-162
2. Kamisaki, Y., Saheki, S., Nakane, M., Palmieri, J.A., Kuno, T., Chang, B.Y., Waldman, S.A., and Murad, F. (1986) Soluble guanylate cyclase from rat lung exists as a heterodimer. *J. Biol. Chem.* **261**, 7236-7241
3. Nakane, M., Arai, K., Saheki, S., Kuno, T., Buechler, W., and Murad, F. (1990) Molecular cloning and expression of cDNAs coding for soluble guanylate cyclase from rat lung. *J. Biol. Chem.* **265**, 16841-16845
4. Suzuki, N., Shimomura, H., Radany, E.W., Ramarao, C.S., Ward, G.E., Bentley, J.K., and Garbers, D.L. (1984) A peptide associated with eggs causes a mobility shift in a major plasma membrane protein of spermatozoa. *J. Biol. Chem.* **259**, 14874-14879
5. Shimomura, H., Dangott, L.J., and Garbers, D.L. (1986) Covalent coupling of a resact analogue to guanylate cyclase. *J. Biol. Chem.* **261**, 15778-15782
6. Garbers, D.L. (1989) Molecular basis of fertilization. *Annu. Rev. Biochem.* **58**, 719-742
7. Singh, S., Lowe, D.G., Thorpe, D.S., Rodriguez, H., Kuang, W.-J., Dangott, L.J., Chinkers, M., Goeddel, D.V., and Garbers, D.L. (1988) Membrane guanylyl cyclase is a cell-surface receptor with homology to protein kinases. *Nature* **334**, 708-712
8. Chinkers, M., Garbers, D.L., Chang, M.-S., Lowe, D.G., Chin, H.M., Goeddel, D.V., and Schulz, S. (1989) A membrane form of guanylate cyclase is an atrial natriuretic peptide receptor. *Nature* **338**, 78-83
9. Schulz, S., Singh, S., Bellet, R.A., Singh, G., Tubb, D.J., Chin, H., and Garbers, D.L. (1989) The primary structure of a plasma membrane guanylate cyclase demonstrates diversity within this new receptor family. *Cell* **58**, 1155-1162
10. Schulz, S., Green, C.K., Yuen, P.S.T., and Garbers, D.L. (1990) Guanylyl cyclase is a heat-stable enterotoxin receptor. *Cell* **63**, 941-948
11. Fülle, H.-J., Vassar, R., Foster, D.C., Yang, R.-B., Axel, R., and Garbers, D.L. (1995) A receptor guanylyl cyclase expressed specifically in olfactory sensory neurons. *Proc. Natl. Acad. Sci. USA* **92**, 3571-3575
12. Yang, R.-B., Foster, D.C., Garbers, D.L., and Fülle, H.-J. (1995) Two membrane forms of guanylyl cyclase found in the eye. *Proc. Natl. Acad. Sci. USA* **92**, 602-606
13. Juilfs, D.M., Fülle, H.-J., Zhao, A.Z., Houslay, M.D., Garbers,

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

D.L., and Beavo, J.A. (1997) A subset of olfactory neurons that selectively express cGMP-stimulated phosphodiesterase (PDE2) and guanylyl cyclase-D define a unique olfactory signal transduction pathway. *Proc. Natl. Acad. Sci. USA* **94**, 3388-3395

14. Schulz, S., Wedel, B.J., Matthews, A., and Garbers, D.L. (1998) The cloning and expression of a new guanylyl cyclase orphan receptor. *J. Biol. Chem.* **273**, 1032-1037
15. Chang, M.-S., Lowe, D.G., Lewis, M., Hellmiss, R., Chen, E., and Goeddel, D.V. (1989) Differential activation by atrial and brain natriuretic peptides of two different receptor guanylate cyclases. *Nature* **341**, 68-72
16. Lowe, D.G., Chang, M.-S., Hellmiss, R., Chen, E., Singh, S., Garbers, D.L., and Goeddel, D.V. (1989) Human atrial natriuretic peptide receptor defines a new paradigm for second messenger signal transduction. *EMBO J.* **8**, 1377-1384
17. Garbers, D.L. and Lowe, D.G. (1994) Guanylyl cyclase receptors. *J. Biol. Chem.* **269**, 30741-30744
18. Laura, R.P., Dizhoor, A.M., and Hurley, J.B. (1996) The membrane guanylyl cyclase, retinal guanylyl cyclase-1, is activated through its intracellular domain. *J. Biol. Chem.* **271**, 11646-11651
19. de Sauvage, F.J., Camerato, T.R., and Goeddel, D.V. (1991) Primary structure and functional expression of the human receptor for *Escherichia coli* heat-stable enterotoxin. *J. Biol. Chem.* **266**, 17912-17918
20. Singh, S., Singh, G., Heim, J.M., and Gerzer, R. (1991) Isolation and expression of a guanylate cyclase-coupled heat stable enterotoxin receptor cDNA from a human colonic cell library. *Biochem. Biophys. Res. Commun.* **179**, 1455-1463
21. Currie, M.G., Fok, K.F., Kato, J., Moore, R.J., Hamra, F.K., Duffin, K.L., and Smith, C.E. (1992) Guanylin: An endogenous activator of intestinal guanylate cyclase. *Proc. Natl. Acad. Sci. USA* **89**, 947-951
22. Hamra, F.K., Forte, L.R., Eber, S.L., Pidhorodeckyj, N.V., Krause, W.J., Freeman, R.H., Chin, D.T., Tompkins, J.A., Fok, K.F., Smith, C.E., Duffin, K.L., Siegel, N.R., and Currie, M.G. (1993) Uroguanylin: Structure and activity of a second endogenous peptide that stimulates intestinal guanylate cyclase. *Proc. Natl. Acad. Sci. USA* **90**, 10464-10468
23. Field, M., Graf, L.H., Jr., Laird, W.J., and Smith, P.L. (1978) Heat-stable enterotoxin of *Escherichia coli*: in vitro effects on guanylate cyclase activity, cyclic GMP concentration, and ion transport in small intestine. *Proc. Natl. Acad. Sci. USA* **75**, 2800-2804
24. Guerrant, R.L., Hughes, J.M., Chang, B., Robertson, D.C., and Murad, F. (1980) Activation of intestinal guanylate cyclase by heat-stable enterotoxin of *Escherichia coli*: Studies of tissue specificity, potential receptors, and intermediates. *J. Infect. Dis.* **142**, 220-228
25. Krause, W.J., Freeman, R.H., and Forte, L.R. (1990) Autoradiographic demonstration of specific binding sites for *E. coli* enterotoxin in various epithelia of the North American opossum. *Cell Tissue Res.* **260**, 387-394
26. Laney, D.W., Jr., Mann, E.A., Dellon, S.C., Perkins, D.R., Giannella, R.A., and Cohen, M.B. (1992) Novel sites for expression of an *Escherichia coli* heat-stable enterotoxin receptor in the developing rat. *Am. J. Physiol.* **263**, G816-G821
27. Laney, D.W., Jr., Bezerra, J.A., Kosiba, J.L., Degen, S.J.F., and Cohen, M.B. (1994) Upregulation of *Escherichia coli* heat-stable enterotoxin receptor in regenerating rat liver. *Am. J. Physiol.* **266**, G899-G906
28. Mann, E.A., Swenson, E.S., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Taguchi, T., Testa, J.R., and Giannella, R.A. (1996) Localization of the guanylyl cyclase C gene to mouse chromosome 6 and human chromosome 12p12. *Genomics* **34**, 265-267
29. Schulz, S., Lopez, M.J., Kuhn, M., and Garbers, D.L. (1997) Disruption of the guanylyl cyclase-C gene leads to a paradoxical phenotype of viable but heat-stable enterotoxin-resistant mice. *J. Clin. Invest.* **100**, 1590-1595
30. Mann, E.A., Jump, M.L., Wu, J., Yee, E., and Giannella, R.A. (1997) Mice lacking the guanylyl cyclase C receptor are resistant to STa-induced intestinal secretion. *Biochem. Biophys. Res. Commun.* **239**, 463-466
31. Seimiya, M., Kusakabe, T., and Suzuki, N. (1997) Primary structure and differential gene expression of three membrane forms of guanylyl cyclase found in the eye of the teleost *Oryzias latipes*. *J. Biol. Chem.* **272**, 23407-23417
32. Yamaguchi, M., Rutledge, L.J., and Garbers, D.L. (1990) The primary structure of the rat guanylyl cyclase A/atrial natriuretic peptide receptor gene. *J. Biol. Chem.* **265**, 20414-20420
33. Takahashi, Y., Nakayama, T., Soma, M., Izumi, Y., and Kanmatsuse, K. (1998) Organization of the human natriuretic peptide receptor A gene. *Biochem. Biophys. Res. Commun.* **246**, 736-739
34. Yang, R.-B., Fülle, H.-J., and Garbers, D.L. (1996) Chromosomal localization and genomic organization of genes encoding guanylyl cyclase receptors expressed in olfactory sensory neurons and retina. *Genomics* **31**, 367-372
35. Perrault, I., Rozet, J.M., Calvas, P., Gerber, S., Camuzat, A., Dollfus, H., Chatelin, S., Souied, E., Ghazi, I., Leowski, C., Bonnemaison, M., Le Paslier, D., Frezal, J., Dufier, J.-L., Pittler, S., Munnich, A., and Kaplan, J. (1996) Retinal-specific guanylyl cyclase gene mutations in Leber's congenital amaurosis. *Nat. Genet.* **14**, 461-464
36. Gigliotti, S., Cavaliere, V., Manzi, A., Tino, A., Graziani, F., and Malva, C. (1993) A membrane guanylate cyclase *Drosophila* homolog gene exhibits maternal and zygotic expression. *Dev. Biol.* **159**, 450-461
37. McNeil, L., Chinkers, M., and Forte, M. (1995) Identification, characterization, and developmental regulation of a receptor guanylyl cyclase expressed during early stages of *Drosophila* development. *J. Biol. Chem.* **270**, 7189-7196
38. Mikami, T., Kusakabe, T., and Suzuki, N. (1998) Molecular cloning of cDNAs and expression of mRNAs encoding α and β subunits of soluble guanylyl cyclase from medaka fish *Oryzias latipes*. *Eur. J. Biochem.* **253**, 42-48
39. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156-159
40. Frohman, M.A., Dush, M.K., and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002
41. Wada, A., Hirayama, T., Kitao, S., Fujisawa, J.-I., Hidaka, Y., and Shimonishi, Y. (1994) Pig intestinal membrane-bound receptor (guanylyl cyclase) for heat-stable enterotoxin: cDNA cloning, functional expression, and characterization. *Microbiol. Immunol.* **38**, 535-541
42. Shimizu, T., Takeda, K., Furuya, H., Hoshino, K., Nomura, K., and Suzuki, N. (1996) A mRNA for membrane form of guanylyl cyclase is expressed exclusively in the testis of the sea urchin *Hemicentrotus pulcherrimus*. *Zool. Sci.* **13**, 285-294
43. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
44. Kozak, M. (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol. Rev.* **47**, 1-45
45. Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132
46. von Heijine, G. (1983) Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* **133**, 17-21
47. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680
48. Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B., Weintraub, H., and Baltimore, D. (1989) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* **58**, 537-544
49. Suh, E., Chen, L., Taylor, J., and Traber, P.G. (1994) A homeodomain protein related to caudal regulates intestine-speci-

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012

fic gene transcription. *Mol. Cell. Biol.* **14**, 7340-7351
50. Arceci, R.J., King, A.A.J., Simon, M.C., Orkin, S.H., and Wilson, D.B. (1993) Mouse GATA-4: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Mol. Cell. Biol.* **13**, 2235-2246
51. Carrithers, S.L., Barber, M.T., Biswas, S., Parkinson, S.J., Park, P.K., Goldstein, S.D., and Waldman, S.A. (1996) Guanylyl cyclase C is a selective marker for metastatic colorectal tumors in human extraintestinal tissues. *Proc. Natl. Acad. Sci. USA* **93**, 14827-14832
52. Gao, Z., Yuen, P.S., and Garbers, D.L. (1997) Interruption of specific guanylyl cyclase signaling pathways in *Advances in Second Messenger and Phosphoprotein Research* (Corbin, J. and Francis, S., eds.) Chapter 15, pp. 183-190, Lippencott-Raven Publishers, Philadelphia
53. Swenson, E.S., Mann, E.A., Jump, M.L., Witte, D.P., and Giannella, R.A. (1996) The guanylin/STa receptor is expressed in crypts and apical epithelium throughout the mouse intestine. *Biochem. Biophys. Res. Commun.* **225**, 1009-1014
54. Yu, S., Avery, L., Baude, E., and Garbers, D.L. (1997) Guanylyl cyclase expression in specific sensory neurons: A new family of chemosensory receptors. *Proc. Natl. Acad. Sci. USA* **94**, 3384-3387

Downloaded from http://jb.oxfordjournals.org/ at Changhua Christian Hospital on October 1, 2012