Diverse Forms of Guanylyl Cyclases in Medaka Fish – Their Genomic Structure and Phylogenetic Relationships to those in Vertebrates and Invertebrates

Authors: Yamagami, Sayaka, and Suzuki, Norio

Source: Zoological Science, 22(8) : 819-835

Published By: Zoological Society of Japan

URL: https://doi.org/10.2108/zsj.22.819
Diverse Forms of Guanylyl Cyclases in Medaka Fish – Their Genomic Structure and Phylogenetic Relationships to those in Vertebrates and Invertebrates

Sayaka Yamagami and Norio Suzuki*

Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan

ABSTRACT—Fish species such as medaka fish, fugu, and zebrafish contain more guanylyl cyclases (GCs) than do mammals. These GCs can be divided into two types: soluble GCs and membrane GCs. The latter are further divided into four subfamilies: (i) natriuretic peptide receptors, (ii) STa/guanylin receptors, (iii) sensory-organ-specific membrane GCs, and (iv) orphan receptors. Phylogenetic analyses of medaka fish GCs, along with those of fugu and zebrafish, suggest that medaka fish is a much closer relative to fugu than to zebrafish. Analyses of nucleotide data available on a web site (http://www.ncbi.nlm.nih.gov/) of GCs from a range of organisms from bacteria to vertebrates suggest that gene duplication, and possibly chromosomal duplication, play important roles in the divergence of GCs. In particular, the membrane GC genes were generated by chromosomal duplication before the divergence of tetrapods and teleosts.

Key words: repetitive interspersed sequence, guanylyl cyclase, medaka fish, linkage group, evolution, genomic structure

INTRODUCTION

By examining the functions of proteins in various organisms, we can often understand how particular molecules or processes help organisms adapt to specific environments. These ingenious survival mechanisms have evolved along with the evolution of organisms. In particular, enzymes play many important roles in cells by using their catalytic potentiality and specificity. Guanylyl cyclases (GCs) are one of the most important enzyme families (EC 4.6.1.2) that catalyze the formation of the second messenger cGMP from GTP.

In 1969, three laboratories identified the enzyme activity that converts GTP to cGMP (Ishikawa et al., 1969; Hardman and Sutherland, 1969; White and Aurbach, 1969; Schultz et al., 1969). As the study of this activity progressed in the mid-1970s, it became relatively certain that the cytoplasmic (soluble type) and particulate-associated (membrane type) forms of GC are present in the homogenates of most of the animal tissues examined (Kimura and Murad, 1974; Chrisman et al., 1975). Soluble GC is a heme-containing heterodimeric protein composed of α and β subunits, and is activated by nitric oxide (NO) and carbon monoxide (CO) (Kamisaki et al., 1986; Drewett and Garbers, 1994; Sharma, 2002). It has been purified to apparent homogeneity from bovine or rat lungs (Koesling et al., 1988, 1990; Nakane et al., 1988, 1990). The soluble GC represents the most important effector enzyme for a signaling molecule NO, which is synthesized by NO synthases (NOS) in a Ca²⁺-dependent manner (Moncada and Higgs, 1995). To date, three isofoms of NOS have been identified and characterized: endothelial NOS (eNOS), inducible NOS (iNOS), and neuronal NOS (nNOS) (Pablo and Garvin, 2003). The binding of NO to the heme leads to a conformation change of soluble GC, resulting in an up to 200-fold increase in the catalytic rate of the enzyme (Koesling and Friebe, 1999). There are two α (α1 and α2) and two β (β1 and β2) subunits of the soluble GC in mammals (Wedel and Garbers, 1997). Many cDNA clones have been isolated from various species (Drewett and Garbers, 1994), including sea urchin (Tanabe and Suzuki, 2001), Drosophila (Shah and Hyde, 1995), and the moth Manduca sexta (Nighorn et al., 1998), which contains the β3 subunit. Each subunit possesses a catalytic domain, but a single subunit itself exhibits no catalytic activ-
ity (Garbers and Lowe, 1994). Soluble GC occurs in relatively high concentrations in vascular smooth muscle cells, platelets, and tissues of the lung, kidney, and brain. Through soluble GC, the NO/cGMP signaling pathway plays important roles in various physiological phenomena, such as vascular smooth muscle relaxation, platelet aggregation, neurotransmission, and programmed cell death (Schmidt and Walter, 1994; Lucas et al., 2000; Fiscus, 2002).

Membrane GC was first purified from its richest source, sea urchin spermatozoa (Garbers, 1976; Radany et al., 1983). It was activated when sea urchin spermatozoa were treated by sea urchin egg peptides called sperm-activating peptides (SAPs) (Suzuki et al., 1981; Hansbrough and Garbers, 1981; Garbers et al., 1982; Suzuki, 1995). In 1988, cDNA for the sea urchin sperm membrane GC was cloned from a testis cDNA library of the sea urchin Arbacia punctulata by Singh et al. Subsequent studies demonstrated that the membrane GC itself is a receptor for SAP-IIa, which is found in A. punctulata egg jelly (Suzuki et al., 1984). On the other hand, it has been known that an atrial natriuretic peptide (ANP) isolated from rat heart is able to activate membrane GC (Waldman et al., 1984), and a partially purified preparation of an ANP-binding membrane protein exhibits GC activity (Kuno et al., 1986; Paul et al., 1987, Takayanagi et al., 1987; Meloche et al., 1988). cDNA of a vertebrate membrane GC was first cloned from a rat brain cDNA library by homology cloning using the sea urchin sperm membrane GC cDNA as a probe; the membrane GC was proved to be the receptor (GC-A) for ANP (Chinkers et al., 1989). This led to the successful cloning of the cDNA for human ANP receptor (GC-A) from a human placenta cDNA library (Lowe et al., 1989). Since then, various membrane GC isoforms have been identified in vertebrates and invertebrates (Drewett and Garbers, 1994; Liu et al., 2004).

In mammals, at least seven membrane enzymes (GC-A through GC-G) have been identified and characterized (Padayatti et al., 2004). They share a topology that consists of an extracellular domain, a transmembrane domain, and an intracellular domain, which is divided further into a kinase-like domain and a catalytic domain (Wedel and Garbers, 1998). It has been suggested that the extracellular domain of all these receptors functions as the specific ligand-binding site, and these enzymes are activated upon ligand binding, although such ligands have been identified for only three membrane GCs (GC-A, GC-B, and GC-C). The other four membrane GCs (GC-D, GC-E, GC-F, and GC-G) are presumed to be orphan receptors (Garbers and Lowe, 1994).

It is known that GC-A is the receptor for atrial natriuretic peptide (ANP) and B-type natriuretic peptide (BNP), which are synthesized by the atrium and ventricles, respectively, and GC-B is the receptor for C-type natriuretic peptide (CNP), which is synthesized mainly in endothelial cells (Chinkers et al., 1989; Lowe et al., 1989; Schulz et al., 1989; Koller et al., 1991; for a review of the NP family Abassi et al., 2004). The GC-A gene has been disrupted in the mouse, which is a good model animal for the study of salt-resistant hypertension (Lopez et al., 1995). It has become evident that GC-A has a local function within the heart or that it influences heart development independent of blood pressure (Oliver et al., 1997), although GC-A has been considered the primary receptor in kidneys and in the smooth muscle vasculature for peptides controlling blood volume, natriuresis, and diuresis. On the other hand, the pathway through GC-B/CNP has an essential role in the local stimulation of growth plate chondrocyte proliferation and differentiation through PKG II, both of which are processes that involve bone formation.

GC-C was first isolated from rat small intestine in 1990, and has been known to be activated by heat-stable enterotoxin (STa) synthesized in Escherichia coli and guanylin, an endogenous ligand in mammals (Schulz et al., 1990). A unique feature of GC-C is that there is an extended carboxyl-terminal tail rich in unchanged polar amino acids, so as to anchor the enzyme to cytoskeletal structures (Schulz et al., 1990). This tail mediates the local effects of STa and guanylin on both intestinal electrolytes and water transport, as well as epithelial cell growth and differentiation (Forst and Cumbe, 1995; Kuhn et al., 1994; Steinbrecher et al., 2002). Additionally, a role of GC-C in liver regeneration has been reported (Schieving and Russell, 1996). Gene disruption studies revealed that GC-C is the physiological receptor for STa, because null mice are resistant to infections by enterotoxigenic bacteria that would otherwise kill wild-type mice (Schulz et al., 1997). Recently, it was demonstrated that fish species such as medaka fish (Oryzias latipes) and eel (Anguilla anguilla) contain two homologues of mammalian GC-C, one of which (OICG9) is activated by STa but not by endogenous ligand(s), and that the transcription of the endogenous ligand-activatable GC (OICG6) is regulated by OIPC4, a medaka fish homologue of mammalian transcriptional positive co-factor (PC4) (reviewed by Nakauchi and Suzuki, 2005).

In mammals, four orphan membrane GCs have been discovered. GC-E and GC-F cDNAs have been isolated from a cDNA library of human retinal photoreceptors (Lowe et al., 1995; Shyjan et al., 1992), and GC-D was cloned from a rat olfactory cDNA library (Fülle et al., 1995). No extracellular ligand has been reported for these three membrane GCs. Instead, two Ca$^{2+}$-sensitive activating proteins (GCAP1 and GCAP2; zebrafish and fugu each contain another, similar protein, GCAP3) appear to bind the kinase-like domains of these GCs and stimulate cGMP formation (Dizhoor and Hurley, 1999; Imanishi et al., 2004). It is understood that these membrane GCs play a key role in the dark-phase recovery of photoreceptor cells (Lowe et al., 1995; Dizhoor et al., 1995; Gorczyca et al., 1995). The other orphan receptor, GC-G, is the last membrane GC to be identified in mammals (Schulz et al., 1998) and was expressed in a wide variety of tissues in the rat. This GC is not activated by any known ligands for membrane GCs.

The homologues of these various mammalian mem-
Table 1. Medaka fish GCs and the corresponding human GCs and ligands*

<table>
<thead>
<tr>
<th>Type of guanylyl cyclases (GCs) (major expressing tissues)</th>
<th>Medaka fish GCs</th>
<th>Corresponding human GCs</th>
<th>Human ligands</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atrial natriuretic peptide receptors (brains, kidney, reproductive organs, and many others)</td>
<td>OlGC1 (89kbp)</td>
<td>hGC-B (16.5kbp)</td>
<td>CNP</td>
</tr>
<tr>
<td></td>
<td>OlGC2 (33kbp)</td>
<td>hGC-A (16.0kbp)</td>
<td>ANP &amp; BNP</td>
</tr>
<tr>
<td></td>
<td>OlGC7 (44kbp)</td>
<td></td>
<td>ANP, BNP</td>
</tr>
<tr>
<td>Enterotoxin/guanylin receptors (intestine)</td>
<td>OlGC6 (15kbp)</td>
<td>hGC-C</td>
<td>Enterotoxin/guanylin</td>
</tr>
<tr>
<td></td>
<td>OlGC9 (?kbp)</td>
<td>Enterotoxin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OlGC8 (13.5kbp)</td>
<td>hGC-G</td>
<td>Enterotoxin</td>
</tr>
</tbody>
</table>

There are no ligands for the following guanylyl cyclases, although in teleosts such as zebrafish and pufferfish, their activities could be stimulated by activating proteins such as GCAP-1, -2, and -3.

<table>
<thead>
<tr>
<th>Species Type</th>
<th>Name of gene</th>
<th>Ensemble gene</th>
<th>GenBank accession No.</th>
<th>Chromosome No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Membrane</td>
<td>Human GC-A</td>
<td>ENSG00000169418</td>
<td>NM_000906</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Human GC-B</td>
<td>ENSG00000159899</td>
<td>NM_003995</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Human GC-C</td>
<td>ENSG00000070019</td>
<td>NM_004963</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Human GC-D</td>
<td>ENSG00000132518</td>
<td>NM_000180</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Human GC-E</td>
<td>AP001189.5.1.186389</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Human GC-F</td>
<td>ENSG00000101890</td>
<td>NM_001522</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Human GC-G</td>
<td>ENSG00000080218</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soluble</td>
<td>Human GCS-α1</td>
<td>ENSG00000164116</td>
<td>U58855</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Human GCS-α2</td>
<td>ENSG00000152402</td>
<td>X63282</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Human GCS-β1</td>
<td>ENSG00000061918</td>
<td>X66534</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Human GCS-β2</td>
<td>ENSG00000123201</td>
<td>NM_004129</td>
<td>13</td>
</tr>
<tr>
<td>Rat Membrane</td>
<td>Rat GC-A</td>
<td>ENSRNOG00000014684</td>
<td>NM_012613</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Rat GC-B</td>
<td>ENSRNOG00000015991</td>
<td>M26896</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Rat GC-C</td>
<td>ENSRNOG00000028106</td>
<td>NM_013170</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Rat GC-D</td>
<td>ENSRNOG00000015058</td>
<td>NM_130737</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rat GC-E</td>
<td>ENSRNOG0000007931</td>
<td>NM_024380</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Rat GC-F</td>
<td>ENSRNOG00000019086</td>
<td>NM_053831</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Rat GC-G</td>
<td>ENSRNOG00000015724</td>
<td>AF024622</td>
<td>1</td>
</tr>
<tr>
<td>Soluble</td>
<td>Rat GCS-α1</td>
<td>ENSRNOG00000012302</td>
<td>U06385</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rat GCS-α2</td>
<td>RNOR0312118.1.59605</td>
<td>AF109963</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Rat GCS-β1</td>
<td>ENSRNOG00000012060</td>
<td>M22562</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rat GCS-β2</td>
<td>ENSRNOG000000094401</td>
<td>NM_012770</td>
<td>15</td>
</tr>
<tr>
<td>Mouse Membrane</td>
<td>Mouse GC-A</td>
<td>ENSMUSG000000027931</td>
<td>L71932</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-B</td>
<td>ENSMUSG000000028469</td>
<td>AL732626</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-C</td>
<td>ENSMUSG000000042638</td>
<td>XM_132928</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-D</td>
<td>ENSMUSG000000030758</td>
<td>AH004877</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-E</td>
<td>ENSMUSG00000020890</td>
<td>NM_008192</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-F</td>
<td>ENSMUSG00000042282</td>
<td>XM_142224</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-G</td>
<td>ENSMUSG00000080218</td>
<td>AK049940</td>
<td>10</td>
</tr>
<tr>
<td>Soluble</td>
<td>Mouse GCS-α1</td>
<td>ENSMUSG00000033910</td>
<td>AF297082</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mouse GCS-α2</td>
<td>ENSMUSG00000041624</td>
<td>XM_150145</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Mouse GCS-β1</td>
<td>ENSMUSG00000028005</td>
<td>AF297083</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mouse GCS-β2</td>
<td>ENSMUSG00000021933</td>
<td>AK047129</td>
<td>14</td>
</tr>
</tbody>
</table>

* Number with kbp in parentheses represents the genomic DNA clone size of the GC.

Table 2. List of the mammalian genes containing the GC regions used for the phylogenetic analyses

<table>
<thead>
<tr>
<th>Species Type</th>
<th>Name of gene</th>
<th>Ensemble gene</th>
<th>GenBank accession No.</th>
<th>Chromosome No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Membrane</td>
<td>Human GC-A</td>
<td>ENSG00000169418</td>
<td>NM_000906</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Human GC-B</td>
<td>ENSG00000159899</td>
<td>NM_003995</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Human GC-C</td>
<td>ENSG00000070019</td>
<td>NM_004963</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Human GC-D</td>
<td>ENSG00000132518</td>
<td>NM_000180</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Human GC-E</td>
<td>AP001189.5.1.186389</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Human GC-F</td>
<td>ENSG00000101890</td>
<td>NM_001522</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Human GC-G</td>
<td>ENSG00000080218</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soluble</td>
<td>Human GCS-α1</td>
<td>ENSG00000164116</td>
<td>U58855</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Human GCS-α2</td>
<td>ENSG00000152402</td>
<td>X63282</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Human GCS-β1</td>
<td>ENSG00000061918</td>
<td>X66534</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Human GCS-β2</td>
<td>ENSG00000123201</td>
<td>NM_004129</td>
<td>13</td>
</tr>
<tr>
<td>Rat Membrane</td>
<td>Rat GC-A</td>
<td>ENSRNOG00000014684</td>
<td>NM_012613</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Rat GC-B</td>
<td>ENSRNOG00000015991</td>
<td>M26896</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Rat GC-C</td>
<td>ENSRNOG00000028106</td>
<td>NM_013170</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Rat GC-D</td>
<td>ENSRNOG00000015058</td>
<td>NM_130737</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rat GC-E</td>
<td>ENSRNOG0000007931</td>
<td>NM_024380</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Rat GC-F</td>
<td>ENSRNOG00000019086</td>
<td>NM_053831</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Rat GC-G</td>
<td>ENSRNOG00000015724</td>
<td>AF024622</td>
<td>1</td>
</tr>
<tr>
<td>Soluble</td>
<td>Rat GCS-α1</td>
<td>ENSRNOG00000012302</td>
<td>U06385</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rat GCS-α2</td>
<td>RNOR0312118.1.59605</td>
<td>AF109963</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Rat GCS-β1</td>
<td>ENSRNOG00000012060</td>
<td>M22562</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rat GCS-β2</td>
<td>ENSRNOG000000094401</td>
<td>NM_012770</td>
<td>15</td>
</tr>
<tr>
<td>Mouse Membrane</td>
<td>Mouse GC-A</td>
<td>ENSMUSG000000027931</td>
<td>L71932</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-B</td>
<td>ENSMUSG000000028469</td>
<td>AL732626</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-C</td>
<td>ENSMUSG00000042638</td>
<td>XM_132928</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-D</td>
<td>ENSMUSG000000030758</td>
<td>AH004877</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-E</td>
<td>ENSMUSG00000020890</td>
<td>NM_008192</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-F</td>
<td>ENSMUSG00000042282</td>
<td>XM_142224</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Mouse GC-G</td>
<td>ENSMUSG00000080218</td>
<td>AK049940</td>
<td>10</td>
</tr>
<tr>
<td>Soluble</td>
<td>Mouse GCS-α1</td>
<td>ENSMUSG00000033910</td>
<td>AF297082</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mouse GCS-α2</td>
<td>ENSMUSG00000041624</td>
<td>XM_150145</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Mouse GCS-β1</td>
<td>ENSMUSG00000028005</td>
<td>AF297083</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mouse GCS-β2</td>
<td>ENSMUSG00000021933</td>
<td>AK047129</td>
<td>14</td>
</tr>
</tbody>
</table>
Fig. 1. (A) A phylogenetic tree of membrane GCs of human, rat, mouse, and medaka fish constructed by using the neighbor-joining method. HpGC denotes *Hemicentrotus pulcherrimus* sperm GC and was used as an outgroup. The number on each node represents a bootstrap value in percent. (B) A phylogenetic tree of soluble GC subunits of human, rat, mouse, and medaka fish constructed by using the neighbor-joining method. The number on each node represents a bootstrap value in percent.
brane GCs in mammals have been identified in the medaka fish *Oryzias latipes*, which is a small freshwater teleost (Kusakabe and Suzuki, 2000), as shown in Table 1. Medaka fish is becoming a very popular model fish, because of recent additions to the genetic toolkit available for this organism, such as BAC, cosmid, and cDNA libraries, an expressed sequence tags (EST) database, genetic linkage mapping, and a whole-genome shotgun sequence approach (Wittbrodt et al., 2002; Shima et al., 2003). Thus far, valuable data have accumulated from genomic studies of puffer fish and from developmental biology studies of zebrafish, and these data are useful to many biologists in performing comparative studies. The genomics of medaka fish is advantageous for such research, because this species' genome contains only 800 million bases, making it half the size of the zebrafish genome and a fourth those of human and mouse (Tanaka, 1995). Medaka fish is also useful in many biological studies because it offers the following benefits: (1) it is a small (3–4 cm) freshwater fish that lays eggs every day and tolerates wide salinity and temperature ranges; (2) under laboratory conditions, its generation time is 6 to 8 weeks, versus 8 to 10 weeks for zebrafish; (3) methods for transgenesis with medaka fish have been developed, and transposon elements have been revealed in the medaka fish genome to use as tools for genetic manipulation (Wakamatsu et al., 1993; Koga et al., 1996, 2002).

### Table 3. List of *C. elegans* genes containing the GC region used for the phylogenetic analyses

<table>
<thead>
<tr>
<th>Type</th>
<th>Name of gene</th>
<th>Ensemble gene</th>
<th>GenBank accession No.</th>
<th>Chromosome No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membrane</td>
<td>gcy-1</td>
<td>AH6.1</td>
<td>NM_063638</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-2</td>
<td>R134.2</td>
<td>NM_063637</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-3</td>
<td>R134.1</td>
<td>NM_063636</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-4</td>
<td>ZK970.5</td>
<td>NM_063817</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-5</td>
<td>ZK970.6</td>
<td>NM_063818</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-6</td>
<td>B0024.6</td>
<td>NM_073249</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-7</td>
<td>F52E1.4</td>
<td>NM_072765</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-8</td>
<td>C49H3.1</td>
<td>NM_068923</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>gcy-9</td>
<td>ZK455.2</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>gcy-10</td>
<td>R01E6.1</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>gcy-11</td>
<td>C30G4.3</td>
<td>NM_078354</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>gcy-12</td>
<td>F08B1.2</td>
<td>NM_062594</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-13</td>
<td>F23H12.6</td>
<td>NM_073696</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-14</td>
<td>ZC412.2</td>
<td>NM_074259</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-15</td>
<td>ZC239.7</td>
<td>NM_062067</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-17</td>
<td>W03F11.2</td>
<td>NM_058542</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>gcy-18</td>
<td>ZK986.8</td>
<td>NM_070048</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>gcy-19</td>
<td>C17F4.6</td>
<td>NM_062090</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-20</td>
<td>F21H7.9</td>
<td>NM_074700</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-21</td>
<td>F22E5.3</td>
<td>NM_061923</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>gcy-22</td>
<td>T03D8.5</td>
<td>NM_075617</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-23</td>
<td>T26C12.4</td>
<td>NM_067908</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>gcy-25</td>
<td>Y105C5B2</td>
<td>NM_070490</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>gcy-27</td>
<td>C06A12.4</td>
<td>NM_070709</td>
<td>IV</td>
</tr>
<tr>
<td></td>
<td>IF75</td>
<td>T01A4.1</td>
<td>NM_058978</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>C04H5.3</td>
<td>C04H5.3</td>
<td></td>
<td>II</td>
</tr>
<tr>
<td></td>
<td>dal-11</td>
<td>B0240.3</td>
<td>NM_073559</td>
<td>V</td>
</tr>
<tr>
<td>Soluble</td>
<td>gcy-31</td>
<td>T07D1.1</td>
<td>NM_076042</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>gcy-32s</td>
<td>C06B3.8</td>
<td>NM_074051</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-33s</td>
<td>F57F5.2</td>
<td>NM_073609</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-34s</td>
<td>M04G12.3</td>
<td>NM_073918</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>gcy-35s</td>
<td>T04D3.4</td>
<td>NM_060943</td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>gcy-36s</td>
<td>C46E1.2</td>
<td>NM_078156</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>gcy-37s</td>
<td>C54E4.3</td>
<td>NM_067770</td>
<td>IV</td>
</tr>
</tbody>
</table>
With the use of these benefits of medaka fish, studies on GCs have advanced significantly in the last decade, and many isoforms of medaka fish GCs have been cloned and characterized. First, OlGC3, OlGC4, and OlGC5 have been isolated from a medaka fish eye cDNA library (Seimiya et al., 1997). Another retina-specific membrane GC was isolated independently from a medaka fish eye cDNA library and designated OlGC-R2 (Hisatomi et al., 1999). Second, two soluble GC subunit cDNA clones have been isolated and characterized, demonstrating that medaka fish has a soluble GC consisting of one α1 subunit and one β1 subunit (Mikami et al., 1998).

Since then, it has been demonstrated that the α1 and β1 subunit genes are aligned in a tandem fashion separating by only a 1 kbp spacer sequence on the medaka fish genome (Mikami et al., 1999). OlGC6 and OlGC9, medaka fish homologues of mammalian GC-C (Mantoku et al., 1999; Iio et al., 2005), and OlGC1, a medaka fish homologue of mammalian GC-B, have been identified, and their complete nucleotide sequences of genomic DNAs have been determined (Mantoku et al., 1999; Takeda and Suzuki, 1999).

In summary, GCs found in medaka fish as well as other vertebrates can be divided into two types of GCs: (1) soluble GCs (α1/β1 and α2/β1 heterodimers) and (2) membrane GCs. The membrane GCs are further divided into four subfamilies: (i) natriuretic peptide receptor, (ii) STa/guanylin receptor, (iii) sensory-organ-specific membrane GC, and (iv) orphan receptor (Table 1).

How many GC members does medaka fish have?

By now, a lot of data have accumulated on the nucleotide sequences of cDNA and genomic DNA of various GCs in various species, and these data are available on the web. Here, we will discuss the relationships between the medaka

![Fig. 2. (A) A phylogenetic tree of membrane GC subunits of Caenorhabditis elegans and medaka fish (OlGC1 and OlGC8) constructed by using the neighbor-joining method. The number on each node represents a bootstrap value in percent. I, II, IX, V, and X denote chromosome number. Gcy-1, gcy-2, and gcy-3 are organized in tandem in the C. elegans genome, as are gcy-4 and gcy-5. (B) A phylogenetic tree of soluble GC subunits of Caenorhabditis elegans, Manduca sexta, and medaka fish constructed by using the neighbor-joining method. The number on each node represents a bootstrap value in percent. The letter ms refers to M. sexta.](https://bioone.org/journals/Zoological-Science/10.2108/zsc2019-000258/fig/fig2.jpg)
fish GCs and those of other species, using the data available on cDNA and genomic DNA clones.

In addition to the two medaka fish homologues (OlGC2 and OlGC7) of mammalian GC-A (Yamagami et al., 2001, 2003) and one medaka fish homologue (OlGC1) of mammalian GC-B (Takeda and Suzuki, 1999), we have isolated another full-length cDNA clone encoding a medaka fish homologue of mammalian STA/guanylin receptor/GC-C and designated it OlGC9 (Iio et al., 2005). Therefore, medaka fish has two different genes, OlGC6 and OlGC9, whose translating products belong to the STA/guanylin receptor/GC-C subfamily. It has been reported that the medaka fish possesses four different sensory-organ-specific membrane GCs: OlGC3, OlGC4, OlGC5, and OlGC-R2 (Seimiya et al., 1997; Hisatomi et al., 1999). Furthermore, a novel membrane GC, OlGC8, was also identified in the medaka fish (Yamagami et al., 2003). OlGC8 possesses a short extracellular domain only half the size of other known membrane GCs, and it was expressed transiently in COS-7 cells that exhibited only basal GC activity. None of the known rat ligands (ANP, BNP, CNP, and C-ANF), medaka fish BNP, medaka fish CNPs (CNP-1, CNP-2, CNP-3, and CNP-4) (Inoue et al., 2003), or various medaka fish tissue extracts stimulated GC activity, suggesting that OlGC8 would be an orphan receptor. On the other hand, to date the cDNA and genomic DNA clones for three soluble GC subunits have been identified in medaka fish and designated OlGCS-α1, OlGCS-β1, and OlGCS-α2, respectively (Mikami et al., 1998; Yao et al., 2003). Therefore, in total, cDNAs and genomic DNAs for 13 GCs, including three soluble GC subunits, have been isolated and characterized in medaka fish (Table 1).

### How many GC members do other species have?

As mentioned above, medaka fish has three NP receptors/membrane GCs (OlGC1, OlGC2, and OlGC7), two STA/guanylin receptors/membrane GCs (OlGC6 and OlGC9), and four sensory-organ-specific membrane GCs (OlGC3, OlGC4, OlGC5, and OlGC-R2). In addition to these, medaka fish contains a unique orphan receptor/membrane GC (OlGC8). On the other hand, it has been reported that mammals possess two NP receptors/membrane GCs (GC-A and GC-B), one STA/guanylin receptor/membrane GC (GC-C), three sensory-organ-specific membrane GCs (GC-D, GC-E, and GC-F), and one orphan receptor/membrane GC (Drewett and Garbers, 1994). It has been reported that fish

<table>
<thead>
<tr>
<th>Linkage groups (LG)</th>
<th>Length (cM)</th>
<th>Medaka fish GC and CNP genes</th>
<th>Human GC genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>44.2</td>
<td>OlGCS-α1/β1</td>
<td>Human GC-A</td>
</tr>
<tr>
<td>2</td>
<td>85.4</td>
<td>OlGC3</td>
<td>Human GCS-α1/β1</td>
</tr>
<tr>
<td>3</td>
<td>84.3</td>
<td>CNP-3</td>
<td>Human GC-B</td>
</tr>
<tr>
<td>4</td>
<td>73.5</td>
<td>OlGC7</td>
<td>Human GC-G</td>
</tr>
<tr>
<td>5</td>
<td>66.0</td>
<td>OlGC1</td>
<td>Human GC-E, Human GCS-α2</td>
</tr>
<tr>
<td>6</td>
<td>82.4</td>
<td>OlGC4, OlGCS-α2</td>
<td>Human GC-C</td>
</tr>
<tr>
<td>7</td>
<td>39.0</td>
<td>CNP-4</td>
<td>Human GCS-β2</td>
</tr>
<tr>
<td>8</td>
<td>57.4</td>
<td>OlGC5</td>
<td>Human GC-D</td>
</tr>
<tr>
<td>9</td>
<td>54.3</td>
<td>OlGCS-α1/β1</td>
<td>X-chromosome: Human GC-F</td>
</tr>
<tr>
<td>10</td>
<td>71.5</td>
<td>OlGCS-α1/β1</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>39.4</td>
<td>OlGC7</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>36.6</td>
<td>OlGC1</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>48.9</td>
<td>OlGC4, OlGCS-α2</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>68.1</td>
<td>OlGC5</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>58.1</td>
<td>OlgC2, CNP-1</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>57.9</td>
<td>OlGC8, OlGC-R2, CNP-2</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>74.7</td>
<td>OlGC6</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>59.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>47.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>71.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>32.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>29.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>47.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>24.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
species such as medaka fish, zebrafish, and fugu contain more members of the same gene family than do mammals (Wittbrodt et al., 1998). Our major question is, How many members of gene families, as well as how many GC genes, do various species have? Recently, the entire shotgun genome sequencing databases of several species have been made publicly available on web sites. To investigate the numbers and varieties of GCs in various animals (human, mouse, rat, zebrafish, fugu, mosquito, fruit fly, Cae-

orhabditis elegans, and C. briggsae), we searched the deduced amino acid sequences of various GC isoforms via the web site of the Ensembl Genome Data Resources, Sanger Institute (http://www.ensembl.org/).

In the human, rat, and mouse databases, we found seven membrane GC isoforms and four soluble GC subunit isoforms in each, (Table 2). Phylogenetic analyses using

### Table 5. List of the numbers of GC genes of various organisms registered in GenBank

<table>
<thead>
<tr>
<th>Species</th>
<th>mGC</th>
<th>sGC</th>
<th>Total</th>
<th>English name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertebrates Mammals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 4 11  human</td>
<td></td>
<td></td>
<td></td>
<td>Homo sapiens</td>
</tr>
<tr>
<td>0 1 1 gorilla</td>
<td></td>
<td></td>
<td></td>
<td>Gorilla gorilla</td>
</tr>
<tr>
<td>0 1 1 chimpanzee</td>
<td></td>
<td></td>
<td></td>
<td>Pan troglodytes</td>
</tr>
<tr>
<td>0 1 1 common gibbon</td>
<td></td>
<td></td>
<td></td>
<td>Hylobates lar</td>
</tr>
<tr>
<td>0 1 1 rhesus monkey</td>
<td></td>
<td></td>
<td></td>
<td>Macaca mulatta</td>
</tr>
<tr>
<td>1 0 1 horse</td>
<td></td>
<td></td>
<td></td>
<td>Equus caballus</td>
</tr>
<tr>
<td>0 1 1 sheep</td>
<td></td>
<td></td>
<td></td>
<td>Ovis aries</td>
</tr>
<tr>
<td>1 1 2 dog</td>
<td></td>
<td></td>
<td></td>
<td>Canis familiaris</td>
</tr>
<tr>
<td>1 0 1 rabbit</td>
<td></td>
<td></td>
<td></td>
<td>Oryctolagus cunicul</td>
</tr>
<tr>
<td>3 1 4 cow</td>
<td></td>
<td></td>
<td></td>
<td>Bos taurus</td>
</tr>
<tr>
<td>1 0 1 pig</td>
<td></td>
<td></td>
<td></td>
<td>Sus scrofa</td>
</tr>
<tr>
<td>2 0 2 domestic guinea pig</td>
<td></td>
<td></td>
<td></td>
<td>Cavia porcellus</td>
</tr>
<tr>
<td>7 4 11 rat</td>
<td></td>
<td></td>
<td></td>
<td>Rattus norvegicus</td>
</tr>
<tr>
<td>7 4 11 house mouse</td>
<td></td>
<td></td>
<td></td>
<td>Mus musculus</td>
</tr>
<tr>
<td>1 0 1 spinifex hopping mouse</td>
<td></td>
<td></td>
<td></td>
<td>Notomys alexis</td>
</tr>
<tr>
<td>Aves</td>
<td></td>
<td></td>
<td></td>
<td>Gallus gallus</td>
</tr>
<tr>
<td>1 0 1 chicken</td>
<td></td>
<td></td>
<td></td>
<td>Xenopus laevis</td>
</tr>
<tr>
<td>3 0 3 clawed toad</td>
<td></td>
<td></td>
<td></td>
<td>Rana catesbeiana</td>
</tr>
<tr>
<td>1 0 1 frog</td>
<td></td>
<td></td>
<td></td>
<td>Bufo marinus</td>
</tr>
<tr>
<td>1 0 1 toad</td>
<td></td>
<td></td>
<td></td>
<td>Ictalurus punctatus</td>
</tr>
<tr>
<td>Teleosts</td>
<td></td>
<td></td>
<td></td>
<td>Oryzias latipes</td>
</tr>
<tr>
<td>10 3 13 medaka fish</td>
<td></td>
<td></td>
<td></td>
<td>Oryzias curvinotus</td>
</tr>
<tr>
<td>1 2 3</td>
<td></td>
<td></td>
<td></td>
<td>Fugu rubripes</td>
</tr>
<tr>
<td>11 6 17 pufferfish</td>
<td></td>
<td></td>
<td></td>
<td>Takifugu rubripes</td>
</tr>
<tr>
<td>0 2 2</td>
<td></td>
<td></td>
<td></td>
<td>Danio rerio</td>
</tr>
<tr>
<td>11 4 15 zebrafish</td>
<td></td>
<td></td>
<td></td>
<td>Pseudopleuronectes americanus</td>
</tr>
<tr>
<td>1 0 1 winter flounder</td>
<td></td>
<td></td>
<td></td>
<td>Anguilla japonica</td>
</tr>
<tr>
<td>2 0 2 Japanese eel</td>
<td></td>
<td></td>
<td></td>
<td>Anguilla anguilla</td>
</tr>
<tr>
<td>2 0 2 European eel</td>
<td></td>
<td></td>
<td></td>
<td>Ictalurus punctatus</td>
</tr>
<tr>
<td>Siluriformes</td>
<td></td>
<td></td>
<td></td>
<td>Squalus acanthias</td>
</tr>
<tr>
<td>1 0 1 channel catfish</td>
<td></td>
<td></td>
<td></td>
<td>Eptatretus cirratus</td>
</tr>
<tr>
<td>Chondrichthyes</td>
<td></td>
<td></td>
<td></td>
<td>Stronglylocentor pusipurus</td>
</tr>
<tr>
<td>1 0 1 spiny dogfish</td>
<td></td>
<td></td>
<td></td>
<td>Asterias amurensis</td>
</tr>
<tr>
<td>Agnatha</td>
<td></td>
<td></td>
<td></td>
<td>Stronglylocentor pusipurus</td>
</tr>
<tr>
<td>1 0 1 hagfish</td>
<td></td>
<td></td>
<td></td>
<td>1 0 1 starfish</td>
</tr>
</tbody>
</table>

Downloaded From: https://bioone.org/journals/Zoological-Science on 14 Feb 2020
Terms of Use: https://bioone.org/terms-of-use
nucleotide sequences of mammalian and medaka fish membrane GC isoforms found no nucleotide sequence homologous to OlGC8 in these genome databases (Table 2 and Fig. 1A). It has been known that humans possess three sensory-organ-specific membrane GC genes (human GC-D, human GC-E, and human GC-F) (Table 2 and Fig. 1A). Phylogenetic analyses indicated that human GC-D forms a clade with rat and mouse GC-E, and that human GC-E

<table>
<thead>
<tr>
<th>Species</th>
<th>mGC</th>
<th>sGC</th>
<th>Total</th>
<th>English name</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Manduca sexta</em></td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>tobacco hornworm</td>
</tr>
<tr>
<td><em>Anopheles gambiae</em></td>
<td>8</td>
<td>4</td>
<td>12</td>
<td>African malaria mosquito</td>
</tr>
<tr>
<td><em>Anopheles aconitus</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Anopheles arabiensis</em></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Anopheles melas</em></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Anopheles quadriannulatus</em></td>
<td>0</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><em>Anopheles minimus C</em></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Anopheles varuna</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td>6</td>
<td>6</td>
<td>12</td>
<td>fruit fly</td>
</tr>
<tr>
<td><em>Apis mellifera</em></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>honey-bee</td>
</tr>
<tr>
<td><em>Bombyx mori</em></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>domestic silkworm</td>
</tr>
<tr>
<td><em>Procambalus clarkii</em></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>red swamp crayfish</td>
</tr>
<tr>
<td><em>Sarcophtes scabiei</em></td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>itch mite</td>
</tr>
<tr>
<td><em>Aplysia sp</em></td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>sea-hares</td>
</tr>
<tr>
<td><em>Limax marginatus</em></td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>slugs</td>
</tr>
<tr>
<td><em>Crassostrea virginica</em></td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>oysters</td>
</tr>
<tr>
<td><em>Caenorhabditis elegans</em></td>
<td>27</td>
<td>7</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td><em>Caenorhabditis briggsae</em></td>
<td>24</td>
<td>7</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td><em>Trichuris vulpis</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Meloidogyne incognita</em></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Heterodera glycines</em></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Paramecium tetraurelia</em></td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tetrahymena pyriformis</em></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Trypanosoma brucei</em></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Plasmodium falciparum</em></td>
<td>7</td>
<td></td>
<td></td>
<td>malaria parasite</td>
</tr>
<tr>
<td><em>Plasmodium yoelii yoelii</em></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Dictyostelium discoideum</em></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Fusarium sporotrichioides</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Magnaporthia grisea</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Agaricus bisporus</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Mycobacterium bovis subsp.</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Vibrio cholerae</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Vibrio vulnificus</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Vibrio parahaemolyticus</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nostoc sp</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nitrosomonas europaea</em></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Mesorhizobium loti</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Prochlorococcus marinus</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Oriza sativa</em></td>
<td>1</td>
<td></td>
<td></td>
<td>rice plant</td>
</tr>
<tr>
<td><em>Arabidopsis thaliana</em></td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>thale cress</td>
</tr>
<tr>
<td><em>Scherffelia dubia</em></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 3. Comparison of the genomic DNA sizes of various medaka fish GCs. Three genes (OlGC1, OlGC2, and OlGC7) are NP receptor-type GCs. Four genes (OlGC3, OlGC4, OlGC5, and OlGC-R2) are sensory-organ-specific GCs. OlGC6 is a gene for an enterotoxin/guanylin receptor-type GC, and OlGC8 is a gene for an orphan receptor-type GC.

Fig. 4. Comparison of the genomic DNA sizes of medaka fish membrane GCs and their homologous genes of fugu and zebrafish. Three genes (OlGC1, OlGC2, and OlGC7) are NP receptor-type GCs. Four genes (OlGC3, OlGC4, OlGC5, and OlGC-R2) are sensory-organ-specific GCs. OlGC6 is an enterotoxin/guanylin receptor-type GC, and OlGC8 is an orphan receptor-type GC. The red, green, and blue boxes indicate the GC genes of medaka fish, fugu, and zebrafish, respectively. Homologues of OlGC5 and OlGC8 were not found in the zebrafish genome databases.
forms a clade with rat and mouse GC-D. OIGC4 forms a clade with human GC-E and rat and mouse GC-D, whereas OIGC3, OIGC5, and OIGC-R2 were not match up in a one-to-one correspondence with mammalian GC-E or GC-F. As shown in Fig. 1B, all mammals searched had four soluble GC subunit genes, although no medaka fish homologue of the mammalian soluble GC β2 subunit gene was identified.

We found 11 membrane GC isoform genes in both the zebrafish and fugu databases, as well as 4 and 6 soluble GC subunit genes in the former and latter, respectively. Since most of the nucleotide sequence data in the database were scattered as either short sequences or sequences with nucleotides missing, it was not easy to align the corresponding amino acids of the respective GCs. Therefore, we carried out phylogenetic analyses using the amino acid sequences of the extracellular domains in the case of zebrafish and medaka fish NP receptors as well as those of a part of the catalytic domain in the cases of the STα/guanylin receptors (GC-C and OIGC6), orphan receptors (GC-G and OIGC8), and various soluble GC subunits, including Manduca sexta GCS-β3 (maGCS-β3). The analyses demonstrated that fugu has all of the genes found in medaka fish, one clone homologous to rat GC-G, and two clones homologous to maGCS-β3. The analyses also revealed that zebrafish contains the same numbers of genes for NP receptors, STα/guanylin receptors, and soluble GC subunits as medaka fish, as well as three sensory-organ-specific receptors/membrane GCs, whereas no clone corresponding to maGCS-β3, rat GC-G, or OIGC8 was found. These findings suggest that medaka fish is much more closely related to fugu than to zebrafish (Wittbrodt et al., 2002), and the genes found in the medaka fish genome have some resemblances to those found in the fugu genome. Assuming that these relationships are the case the medaka fish would have two more membrane GC genes (one that forms a clade with OIGC8 and another that forms a clade with rat GC-G) and three more soluble GC subunit genes (one that forms a clade with rat soluble GC β2 and the other two that each form a clade with maGCS-β3).

We found eight or six membrane GC isoform genes and four or six soluble GC subunit genes in the genome databases of mosquito or fruit fly, respectively. The phylogenetic analyses were executed with five isolated GC clones of M. sexta, including maGCS-β3, suggesting that the M. sexta membrane GCs can be classified into two groups, NP receptor-like membrane GCs and others, although the insect does not seem to possess a sensory-organ-specific membrane GC or an STα/guanylin receptor-like membrane GC. In the phylogenetic tree of M. sexta GCs, the soluble GC subunits each form a clade at the independent position,
suggesting that the evolutionary origin of the soluble GC in insects is different from that of soluble GCs in mammals or fishes and might be functionally different as well.

On the other hand, 27 membrane GC isoform genes and 7 soluble GC subunit genes were found in the database of the nematode *C. elegans* (Table 3). These were subjected to phylogenetic analyses with two medaka fish membrane GCs (OlGC1 and OlGC8) (Fig. 2A). The results indicated that several isoforms located on the same chromosome together form a clade. Therefore, we presume that these isoforms occurred by tandem gene duplications.

Linkage analyses of the medaka fish genes provided 24 linkage gene groups (LGs) with which to compare the arrangements of the orthologous genes among different species (Naruse et al., 2000). Eight membrane GC genes (OlGC1 to OlGC7 and OlGC-R2) are located on separate LGs, that is, on separate chromosomes (Naruse et al., 2000) (Table 4). The OlGC-R2 and OlGC8 genes are located on the same chromosome (LG18) but in positions distant from each other. The OlGCS-α1 and OlGCS-β1 genes are aligned tandemly in LG1, which is different from the location of the OlGCS-α2 gene (LG13). Similarly, the human soluble GC subunit genes (*human GCS-α1, β1*) are located on the same chromosome as each other (LG4) but the chromosome number is different from that of *human GCS-α2* (LG11) (Table 4) (Lucas et al., 2000). A comparison of the location of the medaka fish GC genes with that of the human GC genes indicated that the OlGC4 and OlGCS-α2 genes are both located on LG13, whereas the *human GC-E* and *human GCS-α2* genes are both on LG11. Phylogenetic analysis of mammalian and medaka fish membrane GCs indicates that both OlGC4 and *human GC-E* are homologues of rat *GC-D* (Fig. 1 and Table 2). This suggests that these membrane GC genes were generated by chromosomal duplication before the divergence of tetrapods and teleosts.

In conclusion, we thoroughly searched the NCBI web site (http://www.ncbi.nlm.nih.gov/) for all GC isoforms (Table 5) and found that data on a large number of GCs are available at this web site for a range of organisms from bacteria to vertebrates, although most of the GCs were fragmental, including some that were only for the extracellular domain and others only for the catalytic domain. However, there were several valuable findings. For example, a new class of

![Fig. 6. Schematic drawing of various medaka fish GC genomic organizations with the inserted repeated nucleotide sequences. Closed boxes indicate exons, and lines indicate introns and flanking nucleotide sequences. The two ellipses in the OlGC2 gene denote the repeated sequences. The colored arrowheads pointing left or right indicate a repeated sequence in the sense or antisense direction, respectively.](https://bioone.org/journals/Zoological-Science)
Genomic Structures of Medaka Fish GCs

GCs, termed AtGCs, which were the first functional plant GCs cloned from Arabidopsis thaliana (Ludidi and Gehring, 2003). We are convinced that, during the evolution of organisms from common ancestors, the duplications of GC genes and probably the chromosomal duplication actually occurred many times. We are also convinced that, in these ancestral organisms, both membrane-bound and soluble type GCs play an important role in cells as producers of a second messenger, cGMP, under various regulations.

Comparative study of genomic clone sizes of medaka fish GCs

To date, the 12 genomic clones of medaka fish GCs have been isolated and their sizes and intron/exon organizations have been characterized (Fig. 3 and Table 1). Some GC genes are extremely large, e.g. about 90 kbp for OIGC1 (an NP receptor/membrane GC gene), about 78 kbp for OIGC3 (a sensory-organ-specific membrane GC gene), and about 41 kbp for OIGCS-α2 (a soluble GC subunit gene). A search of a web site of Ensembl Genome Data Resources, Sanger Institute, for data on GCs found GC genes in the fugu genome databases, and several medaka fish GC genes are found in the zebrafish genome databases. A comparison of the membrane GCs’ genomic DNA sizes between medaka fish, fugu, and zebrafish demonstrated that the genomic DNAs of zebrafish membrane GCs corresponding to OIGC2, OIGC7, OIGC3, and OIGC-R2 are twice as large as those of medaka fish (Fig. 4), while those of fugu membrane GCs are less than half the size of those of medaka fish membrane GCs, except for OIGC2. This indicates that the relative size of the genomic DNA of a membrane GC is almost in proportion to the size of the whole genome; that is, the genome of medaka fish is twice as large as that of fugu and half as large as that of zebrafish.

The renovatable search for repetitive sequences in the introns of NP receptors

In a previous study (Yamagami et al., 2001) we

Fig. 7. Schematic drawing of genomic organizations of three NP receptors/GCs with inserted repeated sequences found in various other genes of medaka fish. Closed boxes indicate exons, and lines indicate introns and flanking nucleotide sequences. The two ellipses in the OIGC2 gene denote the repeated sequences. The colored arrowheads pointing left or right indicate a repeated sequence in the sense or antisense direction, respectively. The GenBank/EMBL/DDBJ accession numbers used for the nucleotide sequence alignments are as follows: DMRT1Y (AY129241), Tp53 (AF212997), DMRT1a (AY157712), PSMB9 (AB073378), AQP (AB083078), Transferrin (D64033), PET (polypeptide elongation factor 1 alpha, AB020734), Tyrosinase (AB010101), GnRH-R1 (AB057677), GnRH-R2 (AB057676), mdGnRH (AB074499), Estrogen receptor (AB033491), ZPC2 (AF331672), ZPC4 (AF331674), ZPC5 (AF331675), Globin (AB083077).
described that OlGC1, OIgC2, and OlGC7 each contained several repetitive sequences or conserved elements in the introns. Most of these sequences and elements were found exclusively in NP receptors/membrane GCs, and only six elements were found in six other genes of medaka fish. Recently, more genomic clones of medaka fish have been registered in the DDBJ/EMBL/GenBank databases, including GC genes such as OlgC3, OlgC4, OlgC5, OlgC-R2, and OlgC5-α2, whose exon/intron organizations have been determined in our laboratory (Fig. 5). Therefore, we carried out the same searches and analyses that we did in 2001 on the medaka fish GC genes using the NCBI BLAST search (http://www.ncbi.nlm.nih.gov/BLAST).

To our astonishment, these repetitive elements were found in more variations in many genomic clones of medaka fish (Figs. 6, 7, and 8). These genomic clones were DMRT1Y, Tp53, DMRT1a, PSMB9, AQP, transferrin, polypeptide elongation factor 1 alpha (PET), tyrosinase, GnRH-R1, GnRH-R2, mdGnRH, estrogen receptor, ZPC2, ZPC4, ZPC5, and globin, as shown in Fig. 7; and MHC class I region (sections 1/2 and 2/2, as well as plscr1, zic4, and zic1 genomic genes, as shown in Fig. 8. As shown in Figs. 6, 7, and 8, the search results indicate that these elements accumulate not only in OlGC1, OlgC2, and OlGC7, but are also found in other medaka fish genes (Naruse et al., 1992). In Figs. 6, 7, and 8, highly repetitive, GC-containing sequences found in the medaka fish genomic clones are indicated by one sky-blue arrowhead and one purple-striped arrowhead. Two retrotransposons, Rex3 (Volff et al., 2001a) and Rex6 (Volff et al., 2001b), were found in intron 6 of OlGC1. Rex6 was found not only in OlGC1 but also in OlgC5-α2, OlgC3, and the MHC class I genomic region (Matsuo et al., 2002), as indicated by flesh-colored rectangle-arrowhead combinations. In these figures, the newly found highly repetitive sequences are indicated by one sky-blue and one purple-striped arrowhead (Fig. 6), green and orange (Fig. 7), or purple and orange (Fig. 8). Surprisingly, these highly repetitive elements were found only in the medaka fish genomes and not in any kind of genes, including the GC genes in other species. This suggests that these elements were inserted into introns of medaka fish GCs after the divergence of fish and therefore will be useful for examining the phylogenetic relationships among fishes.

Fig. 8. Schematic drawing of genomic organizations of three NP receptors/GCs with the inserted repeated sequences found in two other medaka fish genes. Closed boxes indicate exons, and lines indicate introns and flanking nucleotide sequences. The two ellipses in the OlGC2 gene denote the repeated sequences. The colored arrowheads pointing left or right indicate a repeated sequence in the sense or antisense direction, respectively. The GenBank/EMBL/DDBJ accession numbers used for the nucleotide sequence alignments are as follows: MHC Class I Region (section 1/2) (AB073376), MHC Class I Region (section 2/2) (AB073377), plscr1, zic4, zic1 genes (AB102768).
Comparative study of physiological role of NP receptors/ GCs of medaka fish and mammals or other species

In mammals, studies of the physiological roles of cGMP signaling, including that of knockout mice of NP receptors, have provided important detail about ligand specificities, salt-resistant hypertension, and bone formation. However, although medaka fish has become a good model animal for the study of the physiological roles of cGMP signaling, studies of ligand specificity and genomic structures have not advanced much thus far. Therefore, we carried out a comparative study of ligand specificity, genomic structure, and genome mapping of NP to surmise the role of medaka NP receptors by comparing the data for mammals. We also analyzed the number of GCs in two other fish species - fugu and zebrafish - as well as in organisms such as the fruit fly and C. elegans, whose whole genome databases are available. One of the important results obtained by the present analyses was that NP receptors existed in all of these species. This suggests that the NP receptor has important roles in various species universally. Above all, the NP family has been differentiated during fish evolution, probably reflecting the changes in osmoregulatory systems promoting adaptation to various osmotic environments. In particular, the NP receptors of the medaka fish have specific roles with several ligands and receptors.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan (no. 11236202) and from the National Project on Protein Structural and Functional Analyses.

REFERENCES


Singh S, Lowe DG, Thorpe DS, Rodriguez H, Kuang WJ, Dangott


(Received July 19, 2005 / Invited Review)