1   *Original research article*

2   **Evolution of the glucagon-like system across fish**

3   João CR Cardoso*+, Rute C Félix+, Carina Costa, Pedro FS Palma, Adelino VM Canário &

4   Deborah M Power*

5

6   Comparative Endocrinology and Integrative Biology, Centre of Marine Sciences, Universidade

7   do Algarve, Campus de Gambelas, 8005-139 Faro, Portugal

8

9   Running title: Fish glucagon-like system

10

11   * Corresponding authors

12   + these authors contributed equally to the work

13   Email addresses:

14   JCRC jccardo@ualg.pt

15   RCF rcfelix@ualg.pt

16   CC carinavcosta@gmail.com

17   PFSP pfpalma@ualg.pt

18   AVMC acanario@ualg.pt

19   DMP dpower@ualg.pt

20

21 **Abstract**

22 In fishes, including the jawless lampreys, the most ancient lineage of extant vertebrates, plasma

23 glucose levels are highly variable and regulation is more relaxed than in mammals. The

24 regulation of glucose and lipid in fishes in common with mammals involves members of the

25 glucagon (GCG)-like family of gastrointestinal peptides. In mammals, four peptides GCG,

26 glucagon-like peptide 1 and 2 (GLP1 and GLP2) and glucose-dependent insulinotropic peptide

27 (GIP) that activate four specific receptors exist. However, in lamprey and other fishes the

28 glucagon-like family evolved differently and they retained additional gene family members

29 (glucagon-related peptide, *gcrp* and its receptor, *gcrpr*) that are absent from mammals. In the

30 present study, we analysed the evolution of the glucagon-like system in fish and characterized

31 gene expression of the family members in the European sea bass (*Dicentrarchus labrax*) a

32 teleost fish. Phylogenetic analysis revealed that multiple receptors and peptides of the

33 glucagon-like family emerged early during the vertebrate radiation and evolved via lineage

34 specific events. Synteny analysis suggested that family member gene loss is likely to be the

35 result of a single gene deletion event. Lamprey was the only fish where a putative *glp1r*

36 persisted and the presence of the receptor gene in the genomes of the elephant shark and

37 coelacanth remains unresolved. In the coelacanth and elephant shark, unique *proglucagon*

38 genes were acquired which in the former only encoded Gcg and Glp2 and in the latter, shared

39 a similar structure to the teleost *proglucagon* gene but possessed an extra exon coding for Glp-

40 like peptide that was most similar to Glp2. The variable tissue distribution of the gene

41 transcripts encoding the ligands and receptors of the glucagon-like system in an advanced

42 teleost, the European sea bass, suggested that, as occurs in mammals, they have acquired

43 distinct functions. Statistically significant ($p < 0.05$) down-regulation of teleost *proglucagon a*

44 in sea bass with modified plasma glucose levels confirmed the link between these peptides and

45 metabolism. The tissue distribution of members of the glucagon-like system in sea bass and

46 human suggests that evolution of the brain-gut-peptide regulatory loop diverged between

47 teleosts and mammals despite the overall conservation and similarity of glucagon-like family

48 members.

49

50 **Keywords:** Glucagon peptide and receptor family, fish, evolution, lineage-specific deletions

51 and duplications, glucose regulation

52

53

Highlights

- The glucagon-like system is present in lamprey and orthologues of *gcgr* and *gipr* are absent
- The glucagon-like system in teleost fish underwent diverse evolutionary trajectories
- The ray-finned fishes lost *glp1r* but the gene was retained in the lamprey
- A single gene encoding *proglucagon* occurs in the coelacanth, spotted gar and elephant shark
- *Proglucagon a* transcripts are modulated by plasma glucose levels in sea bass

1. Introduction

In mammals, energy homeostasis and metabolism is controlled by a series of brain-gut peptides, among them the glucagon-like family. Members of this family include the peptides glucagon (GCG), glucagon-like peptide 1 (GLP1), glucagon-like peptide 2 (GLP2) that are encoded by the proglucagon transcript and glucose-dependent insulinotropic polypeptide (GIP) encoded by *GIP*. The glucagon-like family are members of the Secretin-family of hormones that have a diversity of physiological roles in metazoans (Campbell and Scanes, 1992; Hoyle, 1998; Sherwood et al., 2000; Irwin, 2001, 2002; Irwin and Zhang, 2006; Cardoso et al., 2010; Ng et al., 2010; Wang et al., 2012). In mammals, proglucagon transcripts are produced primarily in the intestine and endocrine pancreas where selective expression and tissue specific proteolytic cleavage occurs. GCG is a 29-amino acid (aa) peptide hormone produced by the pancreatic α-cells (Mojsov et al., 1986; Kieffer and Habener, 1999) which counteracts insulins action on glucose and lipid metabolism (Jiang and Zhang, 2003; Ramnanan et al., 2011). GLP1 and GLP2 derive from the proglucagon precursor and are mainly liberated by the intestinal L cells after nutrient ingestion (Lund et al., 1982; Bell, 1986). The biologically active 30 aa GLP1 (7-37) (Holst, 2007) is an intestinal incretin hormone in mammals that is rapidly released into the circulation to stimulate insulin secretion (Drucker et al., 1987; Holst et al., 1987; Mojsov et al., 1987) and glucose uptake by the liver (Holst et al., 1987; Meier and Nauck, 2005; Baggio and Drucker, 2007). GLP1 is also involved in the inhibition of gastric emptying (Wettergren et al., 1993), regulation of food intake (Donahey et al., 1998) and the secretion of GCG and somatostatin (D'Alessio et al., 1989; Komatsu et al., 1989). GLP2 is a 33 aa peptide, which promotes food absorption and metabolism. It is an important intestinotrophic factor and administration of GLP2 to nude mice (*Mus musculus*) increased small bowel weight and nutrient and energy absorption (Drucker et al., 1996; Estall and Drucker, 2006). GIP, the only product of the *GIP* gene, is a 42 aa peptide produced in the submandibular salivary glands, in the intestine by intestinal K cells and is liberated into the stomach (Buchan et al., 1978; Takeda et al., 1987; Tseng et al., 1993; Yeung et al., 1999). GIP is an incretin peptide (Pederson et al., 1975; Ross and Dupre, 1978; Seino et al., 2010; Wang et al., 2012), an anabolic hormone for adipocyte lipid metabolism, and stimulates lipoprotein lipase activity and fat storage (Eckel et al., 1979).

Four receptors for the glucagon-like peptides exist in mammals. These receptors belong to the Secretin-G protein coupled receptor (GPCR) family and are characterized by seven transmembrane (TM) domains and a long N-terminal region with six conserved cysteine

96   residues important for peptide binding and receptor activation and signalling (Harmar, 2001;
97   Lagerstrom and Schioth, 2008; Bortolato et al., 2014). In contrast to the tissue specific
98   expression of the glucagon-like peptide precursors in humans and rodents, the receptors have
99   a widespread tissue distribution and a diversity of biological functions in vertebrates (Campos
100   et al., 1994; Christophe, 1996; Seino et al., 2010; Pyke et al., 2014).

101   In fish homologues of the mammalian glucagon-like peptides and receptors have been
102   reported. Gcg and Glp-1 peptides were isolated from lamprey intestinal tissue and the teleost
103   pancreas and intestine (Table 1) (Conlon et al., 1991; Conlon et al., 1993a; Conlon et al.,
104   1993b), but so far, Glp-2 has never been isolated and purified and its function is uncertain. In
105   lamprey, the peptides derive from two unique *proglucagon* genes (*proglucagon I* and *II*) that
106   originated from a lineage specific event (Irwin et al., 1999; Wang et al., 1999; Irwin, 2001). In
107   the teleosts, specific genome tetraploidization means duplicate gene family members exist
108   (Cardoso et al., 2005; Roch et al., 2009; Ng et al., 2010; Hwang et al., 2014; Irwin, 2014). In
109   teleosts, two *proglucagon* genes with distinct coding potential generate *proglucagon a* and *b*
110   (Table 1). *Proglucagon a* encodes Gcg, Glp1 and Glp2 and is mainly expressed in the intestine
111   and *proglucagon b* encodes Gcg and Glp1, but not Glp2 and is mainly expressed in the pancreas
112   (Irwin and Wong, 1995; Plisetskaya and Mommsen, 1996; Zhou and Irwin, 2004). The piscine
113   Gcg peptide is primarily produced by the α-cells of the pancreatic islets however expression
114   has also been observed in the L-cells of the intestine (Plisetskaya and Mommsen, 1996). It
115   shares high sequence similarity with mammalian GCG and also has a similar function acting
116   as a potent antagonist of insulin (Navarro and Gutierrez, 1995; Navarro et al., 2002). Gcg
117   significantly increases glucose production in goldfish (*Carassius auratus*) and rainbow trout
118   (*Oncorhynchus mykiss*) hepatocytes (Harmon and Sheridan, 1992; Chow et al., 2004) and in
119   the gilthead sea bream (*Sparus aurata*) it acts on adipose tissue and is a potent lipolytic
120   hormone (Albalat et al., 2005). The GLP1 peptide sequence is generally poorly conserved
121   across fish, and unlike the mammalian homologue it is not an incretin hormone, it does not
122   stimulate insulin release and has similar actions to Gcg (Mommsen et al., 1987; Mommsen and
123   Moon, 1989; Plisetskaya et al., 1989; Plisetskaya and Mommsen, 1996; Mojsov, 2000; Polakof
124   et al., 2011a). Injections of human GLP1 in fish inhibits food intake in the channel catfish
125   *Ictalurus punctatus* (Silverstein et al., 2001; Volkoff et al., 2005). A single Gip peptide
126   precursor occurs in fish and mammals in the intestine and pancreas but its functions remain
127   poorly explored (Musson et al., 2009; Musson et al., 2011).

128   We previously reported the evolution of the Secretin-family system and its functional
129   specialization in the context of the mechanisms underpinning the retention of duplicate genes
130   in teleosts (Guerreiro et al., 2007; Cardoso et al., 2014a; Cardoso et al., 2014b; Martins et al.,
131   2014; Cardoso et al., 2016). Glucagon-like receptors occur in fish (Cardoso et al., 2010; Ng et
132   al., 2010; Hwang et al., 2013; Irwin, 2014) and in teleosts duplicate *gcgr* (*gcgra* and *gcgrb*)
133   genes exist, the gene homologue of *glp1r* is absent and *gipr* has been described in relatively
134   few species. Recently, a novel *gcg*-related peptide (*gcrp*) and its cognate receptor was
135   described in zebrafish, which is absent in mammals (Wang et al., 2012; Park et al., 2013). In
136   the present study taking into consideration the well described functional role of *gcg* family
137   members in studies from the 1980's and 1990's but the paucity of studies placing them in an
138   evolutionary context we characterized glucagon-like receptor and peptide evolution in the
139   extant fishes, the most successful group of vertebrates (Venkatesh, 2003; Ravi and Venkatesh,
140   2008). Glucagon-like receptors and peptides were identified in the genomes of several fish
141   including the lampreys (*Petromyzon marinus* and *Lethenteron japonicum*) (Mehta et al., 2013;
142   Smith et al., 2013), the cartilaginous elephant shark (*Callorhinchus milii*) (Venkatesh et al.,
143   2014), the freshwater ray-finned spotted gar (*Lepisosteus oculatus*) (Braasch et al., 2016), the
144   lobe-finned fish coelacanth (*Latimeria chalumnae*) (Amemiya et al., 2013) and several teleosts.
145   The evolution of the glucagon-like family in fish is placed in the context of the evolution of
146   the gastro-entero-pancreatic system.
147
148   **2. Methods**
149   *Sequence database searches*
150   Glucagon-like receptor genes and peptide gene precursors were procured in several fish
151   genomes using tBLASTN (January 2017). Searches for the receptors were performed using the
152   zebrafish (*Danio rerio,* Gcgra, XP_693606, Gcgrb XP_691434, Glp2r XP_009304634 and
153   Gipr XP_005157796*),* medaka Gcrpr (*Oryzias latipes*, ENSORLP00000009481) and human
154   GLP1R (*Homo sapiens*, NP_002053.3) as bait. Searches for genes encoding glucagon-like
155   peptides were performed using the deduced mature peptides of zebrafish proglucagon
156   (proglucagon a, NP_001258699 and proglucagon b, NP_001229699) and the mature Gip
157   (NP_001073528) and chicken Gcrp (*Gallus gallus,* XP_015155872.1) peptides as bait. The
158   fish genomes explored included the lobe-finned coelacanth (*Latimeria chalumnae*,
159   http://www.ensembl.org), 13 ray-finned fish genomes (12 teleosts and the primitive freshwater
160   ray-finned fish the spotted gar, *Lepisosteus oculatus* available from http://www.ensembl.org),

161   the Chondrichthian elephant shark (*Callorhinchus milii*, http://esharkgenome.imcb.a-
162   star.edu.sg) and two Jawless fish species the Sea lamprey (*Petromyzon marinus*,
163   http://www.ensembl.org) and the Japanese lamprey (*Lethenteron japonicum*,
164   http://jlampreygenome.imcb.a-star.edu.sg). The teleost fish included 8 species available from
165   http://www.ensembl.org (2 pufferfishes *Tetraodon nigroviridis* and *Takifugu rubripes*;
166   stickleback, *Gasterosteus aculeatus*; tilapia, *Oreochromis niloticus*; medaka, *Oryzias latipes*;
167   platyfish, *Xiphophorus maculatus*; Atlantic cod, *Gadus morhua* and cavefish, *Astyanax*
168   *mexicanus*), the genomes of the European sea bass (*Dicentrachus labrax*, http://public-
169   genomes-ngs.molgen.mpg.de (Tine et al., 2014)), the Japanese eel (*Anguilla japonica*) from
170   http://www.zfgenomics.org/sub/eel (Henkel et al., 2012), the smooth tongue sole (*Cynoglossus*
171   *semilaevis* (Chen et al., 2014) retrieved from the NCBI databases and Atlantic salmon (*Salmon*
172   *salar* (Davidson et al., 2010) retrieved from the NCBI and from the species genome assembly
173   (http://salmobase.org/). Sequence hits that shared highest homology for the queries and other
174   sequences annotated *in silico* as members of the glucagon-like system were retrieved and their
175   identity was confirmed against the human homologues. In order to account for potential
176   genome assembly errors and gene absence, the deduced protein sequences of the glucagon-like
177   receptors and putative activating mature peptides were used to search teleost specific EST
178   databases (tax id: 32443, https://www.ncbi.nlm.nih.gov) and sequence hits with an e value < e
179    $^{-20}$ were retrieved and analysed.

180   Members of the glucagon-like gene family were also retrieved from
181   http://www.ensembl.org/ using orthologue gene annotation for the amphibian (*Xenopus*
182   *tropicalis*), Anole lizard (*Anolis carolinensis*), chicken (*Gallus gallus*) and human (*Homo*
183   *sapiens*). The duplicate urochordate *Ciona intestinalis* putative receptors included in the
184   analysis were the same as in Cardoso et al. (2006).

185

186   *Phylogeny and sequence alignments*

187   Phylogenetic analysis of the glucagon-like deduced amino acid sequences of the
188   receptors (105 sequences) and mature peptides (123 sequences) were performed using
189   sequence alignments produced by Muscle software available from Aliview (Larsson, 2014).
190   The mature glucagon-like peptides were manually retrieved from the peptide precursors by
191   localizing putative basic peptide proteolytic consensus cleavage sites or predicted using the
192   Neuropred tool, (http://neuroproteomics.scs.illinois.edu/neuropred.htm, (Southey et al., 2006)
193   and available peptide sequence data. For the receptors only the regions from TM1 to TM7

194   (identified based on annotated human TM predicted available from UniProt,

195   http://www.uniprot.org) including the intra and extracellular loops were used for phylogenetic

196   analysis (Supplementary File 1). For the peptides, the trees were constructed using the highly

197   conserved (1-28 aa) sequence alignment of the deduced Gcg, Glp1, Glp2, Gip and Gcrp mature

198   peptides from fish and tetrapod (Supplementary File 2, Supplementary File 3).

199   The edited receptor and peptide sequence alignments were analysed in ProtTest (2.4)

200   according to the Akaike Information Criterion (AIC) statistical model (Abascal et al., 2005) to

201   select the best model to build the tree. Phylogenetic trees were constructed with the Bayesian

202   inference (BI) method in MrBayes 3.2 (Ronquist et al., 2012) and the Maximum likelihood

203   (ML) method using PhyML (v3.0 aLRT). Construction of receptor and peptide phylogenetic

204   trees was performed using a JTT substitution model (Jones et al., 1992). The receptor BI tree

205   was built sampling 1.000.000 generations and posterior probability values supporting tree

206   branching included. The receptor ML tree was built with a gamma shape (4 rate categories) of

207   G = 0.961 and a proportion of invariable sites of I = 0.028 with 100 bootstraps replicates

208   (Felsenstein, 1985). Receptor trees were rooted on human PTH1R (ENSP00000321999). The

209   phylogenetic trees of the deduced peptides were unrooted and the BI tree was performed

210   sampling 1.000.000 generations and including the posterior probability values to support tree

211   branching. The ML phylogenetic tree was constructed with a gamma shape (4 rate categories)

212   of G = 1.649 using 100 bootstrap replicates. Amino acid sequence identities were calculated

213   using GeneDoc software (http://iubio.bio.indiana.edu/).

214

215   *Gene neighbourhood analysis*

216   To establish how the *gipr* and *gcrpr* and peptide genes evolved across the fishes the

217   genome regions flanking the genes were characterized in the European sea bass and tilapia and

218   the neighbouring gene environment was compared with other teleosts. Medaka was selected as

219   the teleost representative that lacks both *gipr* and *gip* genes and zebrafish as the representative

220   species in which *gcrpr* and *gcrp* are absent. Gene homologues were identified using available

221   genome annotation (Genomicus, http://www.genomicus.biologie.ens.fr, (Louis et al., 2013)

222   and was complemented with sequence homology searches to confirm gene presence or absence.

223   Comparisons were extended to the homologue genome regions in the elephant shark, spotted

224   gar, coelacanth and human. Searches were also performed in lamprey but the relatively short

225   genome fragments impeded the identification of syntenic genome regions and this data was not

226 included. The *glp1r* gene environment in the elephant shark, spotted gar and coelacanth was
227 also characterised and compared to human.
228

229 *Biological material*
230   European sea bass (weight 539 ± 45.47 g, length 33.32 ± 1.02 cm) was maintained at
231 Ramalhete Marine Station (Centre of Marine Sciences, University of Algarve) in 500 L flow-
232 through seawater tanks under natural winter temperatures and normal photoperiod and fed with
233 commercial dry pellets at 1% of body weight/day. Tissue samples (brain, duodenum, adipose
234 tissue, kidney and liver) from three animals were collected and immediately frozen in liquid
235 nitrogen and stored at - 80 ℃ until required for characterization of the tissue distribution of
236 receptor and peptide transcripts.
237   To correlate receptor gene expression with changes of plasma glucose levels, tissue
238 samples from juvenile sea bass (weight 99.53 ± 3.02 g, length 20.23 ± 0.12 cm) that were
239 maintained without feeding for 3 days in 500 L flow-through seawater tanks at 16-17 ℃ and
240 under normal winter photoperiod (December 2012) were collected. Based on plasma glucose
241 levels (Spinreact 1001190 kit, Barcelona, Spain) two experimental groups (n=6 individuals
242 each) were defined: low glucose (4.29 ± 0.48 mmol/L) and high glucose (11.37 ± 1.81 mmol/L)
243 and tissue samples from liver, adipose tissue, brain and duodenum were analysed.
244   All animals were anesthetized with 2-phenoxyetanol (Sigma-Aldrich) before being
245 killed by decapitation. Animal manipulations were performed in compliance with international
246 and national ethics guidelines for animal care and experimentation, under a "Group-I" license
247 from the Portuguese Government Central Veterinary service.
248

249 *RNA extractions and cDNA synthesis*
250   Total RNA (tRNA) was extracted using a Maxwell 16 Total RNA purification system
251 (Promega) and a maximum of 10 μg was treated with 1 U DNase (DNA-free Kit, Ambion) for
252 30 min at 37 °C according to the manufacturer's instructions. DNase treated tRNA (500 ng)
253 was denatured at 65 °C for 5 min, quenched on ice for 5 min and used for cDNA synthesis in
254 a 20 μl reaction volume containing 10 ng of pd(N)6 random hexamers (Jena Bioscience), 2
255 mM dNTPs (Thermoscientific), 5 U of RevertAid (Thermoscientific) and 0.4 U Ribolock
256 RNase inhibitor (Thermoscientific). cDNA was synthesized for 10 min at 20 °C followed by
257 60 min at 42 °C and 5 min at 72 °C and the quality of the synthesis was assessed by evaluating
258 the amplification of *18s* ribosomal subunit (Table 1). The following thermal cycle was used:

259 95 ºC, 3 min; (95 ºC, 10 sec, 58 ºC, 10 sec, 72ºC 10 sec) cycled 25 times and 5 min at 72 ºC

260 and amplified products were visualized on a 2 % Agarose/1xTAE gel stained with Greensafe

261 Premium (NZY Tech, Portugal).

262

263 *Quantitative expression*

264 Expression of the sea bass glucagon-like receptors and precursors was carried out by

265 quantitative real-time PCR (q-PCR) with transcript specific primers (Table 2) in a 10 μl final

266 reaction volume that contained 200 nM of forward and reverse primer, SsoFast EvaGreen

267 supermix (Bio-Rad, Portugal) and 2 μl of template cDNA (diluted 1:5). Elongation factor 1-

268 alpha (*ef1α*) and 18S ribosomal subunit (*18s*) were used as reference genes (cDNA diluted

269 1:100 and 1:1000, respectively) as no significant differences ($p < 0.05$) in transcript abundance

270 were found between the samples (One-way analysis-ANOVA, data not shown). qQ-PCR

271 analysis was performed in duplicate reactions (< 5 % variation between replicates) using a CFX

272 Connect™ Real-Time PCR Detection System for 96-well microplates (Bio-Rad, Portugal).

273 Optimized cycling conditions consisted of 95 ºC for 30 sec, followed by 44 cycles of 95 ºC for

274 5 sec and the appropriate annealing temperature for 10 sec. Melting curves were performed to

275 detect nonspecific products and primer dimers. q-PCR efficiencies and R^2 (coefficient of

276 determination) were established with standard curves prepared in duplicate from a 10-fold

277 serial dilution series of the purified PCR product of each of the target genes. The resulting

278 expression values of the target genes were normalized with the geometric mean of the

279 expression levels of the reference genes. Control reactions (tRNA instead of cDNA) was

280 included to confirm the absence of genomic DNA contamination. The amplicons of q-PCR

281 reactions were sequenced to confirm the specificity of the q-PCR.

282

283 *Statistical analysis*

284 Results are presented as the mean ± SEM. Statistical differences between different

285 experimental groups were detected using an unpaired Student's t-test (two-tail, confidence

286 level 95%). The significance cut-off was taken at p < 0.05. The analysis was performed with

287 Prism GraphPad software (7.0).

288

289 **3. Results**

290 *The glucagon-like peptide system in fish*

291   Searches in fish genomes identified genes for both peptides and receptors of the
292 glucagon-like peptide system in all the species analysed. These included homologues of human
293 and zebrafish receptor and peptide precursor genes and for medaka *gcrpr* and chicken *gcrp*
294 precursor genes (Figure 1).
295
296   *Glucagon-like peptide receptors*
297   At least 3 glucagon-like receptor orthologues of the human and teleost glucagon-like
298 peptide receptors were found in fish. In lampreys, 4 receptor members were retrieved, in the
299 elephant shark and coelacanth 3 and 5 were found (Figure 1). In the spotted gar genome 4
300 receptor genes exist and in the teleosts receptor gene number varied from 3 to 5. In teleosts in
301 general, two *gcgr* genes and a single *glp2r* gene were found and the *gipr* gene was generally
302 present when the *gcrpr* gene was absent. For example, a putative *gipr* gene was present in the
303 Japanese eel, zebrafish, cavefish and Atlantic salmon and the *gcrpr* gene was present in
304 *Tetraodon*, *Takifugu*, stickleback, medaka, smooth tongue sole and Atlantic cod. Of the teleosts
305 analysed the sea bass and tilapia were the only species where genes for both *gipr* and *gcrpr*
306 persisted (Figure 1).
307
308   *Glucagon-like peptide precursors*
309   Genes for fish proglucagon and Gip precursor and the coding exon for the gcrp peptide
310 were identified. In lamprey, the two *proglucagon* genes (*proglucagon I* and *proglucagon II*)
311 were found although the previously reported *gip* (Musson et al., 2011) was not identified in the
312 current genome assembly. In elephant shark, the *proglucagon* gene encoded three putative Glp-
313 like peptides (designated Glp-like a, Glp-like b, and Glp-like c) and sequence alignment
314 revealed that Glp-like b and Glp-like c were more similar with human GLP1 and GLP2,
315 respectively (Figure 2). The newly identified Glp-like a shared highest sequence identity with
316 the elephant shark putative Glp2 (Glp-like *c*, 39% aa identity) than with Glp1 (Glp-like *b*, 28%
317 aa identity). In the spotted gar, a single *proglucagon* gene encoded the three peptides but in
318 coelacanth the gene homologue only exons for Gcg and a single Glp-like peptide (that shared
319 highest sequence identity to human GLP2, 57% aa) were found (Figure 2).
320   In teleosts, with the exception of eel, duplicate *proglucagon* genes (*proglucagon a* and
321 *proglucagon b*) were identified. The deduced sequences of the teleost duplicate mature Gcg
322 and Glp1 peptides were highly identical. The paralogue mature Gcg and Glp1 peptides from
323 sea bass shared 85% (Gcg a and Gcg b) and 75% (Glp1 a and Glp 1b) aa sequence identity.

324   The deduced Glp2 peptide in sea bass shared 89% sequence identity with the zebrafish peptide

325   and 75% identity with Glp2 in the spotted gar but conservation with the mammalian homologue

326   was generally poor (< 50% identity).

327   The *gip* and *gcrp* genes were present in the teleost genomes as was the putative cognate

328   receptor gene (Figure 1). The *gip* gene was absent from the coelacanth and spotted gar but a

329   putative *gcrp* gene was identified and in the elephant shark the *gip* gene persisted and *gcrp* was

330   absent. In tilapia both *gip* and *gcrp* genes persisted and the presence of a *gip* gene was deduced

331   based on the identification of an expressed sequenced tag (GR702825) that contained the

332   complete peptide precursor. Searches against the tilapia genome revealed it mapped to

333   chromosome fragment GL831200.1 (2,08 Mb). Despite the existence of a putative *gipr*, a *gip*

334   gene was not found in the sea bass genome or in databases of transcript sequences. Attempts

335   to amplify it by PCR also failed, indicating it has most likely been lost from the sea bass

336   genome. The deduced Gcrp peptide from tilapia shared 93% aa identity with the sea bass

337   peptide and 72% aa identity with the *Xenopus* homologue. The tilapia Gip mature peptide

338   shared 78% aa identity with the zebrafish homologue and 42% identity with human GIP.

339   Salmon was the only teleost where two *gip* genes (*gip a* and *gip b*) precursors were found

340   (mature peptides are 100% aa identical) presumably a consequence of the salmon specific

341   genome duplication (Macqueen and Johnston, 2014). In the elephant shark the *gip* precursor

342   gene was found and the deduced mature peptide shared 50% and 40% aa identity with the

343   zebrafish and human homologues, respectively.

344   Analysis of the deduced amino acid sequence of the proglucagon and gip precursors

345   revealed that in most of the fish species the consensus dibasic (KR, RR) and monobasic (R)

346   proteolytic cleavage sites flanked the predicted mature peptides. In spotted gar, medaka and

347   stickleback proteolytic cleavage sites were absent from the putative Glp2 C-terminal region

348   and thus the size of these peptides remains to be established. In these cases, the size of the Glp-

349   2 peptide was predicted based on the conserved length (35 aa) of the homologue mature

350   peptides in other ray-finned fish. The sea bass was the only teleost in which a longer Glp2

351   peptide of 42 aa was predicted, although for determination of sequence conservation and

352   phylogenetic analysis only the first 35 aa were considered (Supplementary File 2).

353   Analysis of the predicted gene structure of the region coding the full-length peptide

354   precursor revealed that fish *proglucagon* had a distinct organization mainly due to the existence

355   of two genes with different peptide coding potential (Figure 3). The lamprey *proglucagon I*

356   had 4 coding exons and *proglucagon II* had 3 coding exons. The teleost gene coding

357 *proglucagon a* was composed of 5 exons and *proglucagon b* was organized in 3 exons. The

358 spotted gar *proglucagon* encoding gene had a similar organization to teleost *proglucagon a* and

359 human *proglucagon*. The coelacanth *proglucagon* was composed of 3 and the elephant shark

360 of 7 coding exons, respectively. In all species, the peptides were encoded on consecutive exons

361 suggesting they emerged through exon duplication events (Figure 3). The gene structure of the

362 fish *gip* coding gene was highly conserved across species and the peptide was encoded by a

363 single exon in all species analysed. In zebrafish and elephant shark the gene structure was

364 similar and spanned 4 exons (data not shown). In all fish species analysed only the potential

365 coding exon for *gcrp* was identified and attempts to find a full-length *gcrp* encoding gene

366 failed.

367

368 *Phylogenetic analysis*

369   Phylogenetic trees built using both BI and ML approaches for the glucagon-like

370 receptors (Figure 4 and Supplementary Figure 1) and peptides (Figure 5 and Supplementary

371 Figure 3) had similar topologies. The clustering of the fish receptors (Figure 4, Supplementary

372 Table 1) and mature peptides (Figure 5, Supplementary Table 2) with their homologues from

373 other species suggested that they emerged early and duplicated early in the vertebrate radiation.

374

375   *Glucagon-like receptors*

376   Five main vertebrate receptor clades were identified and the tree topology suggested

377 that the ancestral *glp2r* gene was the first to diverge. The *gcrpr* and the *gipr/gcgr* precursor

378 genes arose subsequently from the same duplication event and members of the fish receptors

379 were distributed in all receptor clades (Figure 4). In lamprey, a receptor orthologue of human

380 GLP1R was found that clustered closely with the tetrapod homologues and clustering of the

381 other lamprey receptor sequences revealed that a *glp2r* gene also existed in addition to

382 duplicates of the gnathostome *gcrpr*. One lamprey *gcrpr* paralogue corresponded to the

383 previously proposed lamprey *gcgr* (Irwin, 2014). Sequence homologues of the gnathostome

384 *gcgr* and *gipr* were not found in both lamprey genomes. In the elephant shark, phylogeny

385 confirmed the identification of *glp2r* and *gcgr* genes but the existence of a *glp1r* was

386 unresolved. This receptor (SINCAMG00000015132) is annotated in the cartilaginous fish

387 genome as a potential human *GLP1R* orthologue and it shares highest sequence similarity for

388 the tetrapod GLP1R (72 % aa sequence). However, receptor clustering in the phylogenetic trees

389 varied according to the method used and only the BI tree grouped the gene in the *glp1r* clade.

390   In the coelacanth, phylogeny clearly identified orthologues of the human GLP2R, GCGR,

391   GIPR and of the non-mammalian *gcrpr* genes and also revealed the existence of a highly

392   divergent glucagon-like receptor gene. The latter gene did not cluster within any of the major

393   receptor clades and its deduced protein sequence shared 68-70% aa similarity with the other

394   glucagon-like receptor members (Figure 4 and Supplementary Figure 1). In the spotted gar a

395   similar gene repertoire to that of the teleosts was found and *glp1r* was also absent. Both ML

396   and BI trees confirmed the duplication of *gcgr* (*gcgra* and *gcgrb*) in teleosts and the identity of

397   the sea bass and tilapia *gipr* and *gcrpr* receptor genes. In eel, two *gcgr* genes were identified

398   that grouped within the teleost *gcgrb* cluster suggesting that they may have arisen from an

399   independent lineage specific gene duplication.

400

401   *Glucagon-like peptides*

402   The fish and other vertebrate glucagon-like mature peptides grouped in 5 main clusters

403   in the phylogenetic tree (Figure 5). The lamprey peptides grouped within the vertebrate Gcg,

404   Glp1 and Glp2 clades indicating that they arose early in evolution and prior to the divergence

405   of cyclostomes and gnathostomes (Figure 5, Supplementary Figure 3). The three elephant shark

406   Glp-like peptides grouped differently: Glp-like b grouped with the vertebrate Glp1 and the

407   remaining two peptides, Glp-like a and Glp-like c grouped with vertebrate Glp2, suggesting

408   that in this species two copies of Glp2 exist. The clustering of the teleost peptides with those

409   of other vertebrates confirmed their identity and the grouping of the salmon Gcg a and b, Glp1a

410   and b and of the two Gip peptides suggested that the *proglucagon* and *gip* precursors are likely

411   to have resulted from the subsequent species-specific genome duplication (Figure 5). The

412   coelacanth Glp-like peptide tended to group within the vertebrate Glp2 cluster.

413

414   *Gene synteny across fish*

415   *Gipr and gcrpr gene neighbourhood*

416   In sea bass *gipr* maps to LG13 and in tilapia to scaffold GL831556 and overall, at least

417   12 genes were identified in synteny with the homologue genome regions in zebrafish

418   chromosome 15 and in medaka chromosome 13, where the gene for this receptor was absent

419   (Figure 6). In sea bass and tilapia, the *gipr* gene was flanked by the *thap8* gene and by the

420   claudins (*cldn*), a family of genes that have expanded in teleosts (Tine et al., 2014) and in

421   medaka both genes were in very close proximity suggesting that the *gipr* gene was deleted from

422   its genome. The gene order of the zebrafish *gipr* gene environment differed from the other

423 teleosts with *gipr* localized near the *eml2* and *gpr4* genes. Comparison of the gene environment

424 in teleosts with that in the spotted gar (LG2) and human (chr 19) revealed that gene order has

425 been conserved with the zebrafish (chr 15) while most of the genes that flank the sea bass and

426 tilapia *gipr* were localized in different genome regions in the spotted gar LG22 and human chr

427 7 (not shown). This suggests that insertion of a gene block occurred within the teleost *gipr*

428 genome region and that, during the teleost radiation gene shuffling also occurred (Figure 6).

429   In sea bass, the *gcrpr* gene maps to chromosome LG1B and in tilapia to scaffold

430 GL831136. At least 7 conserved neighbouring genes were found for *gcrpr* bearing species

431 (medaka chromosome 19 and spotted gar chromosome LG13) and in zebrafish (chromosome

432 24) in which this gene was absent (Figure 6). The well-conserved genome region found when

433 the zebrafish genome was compared with the homologue region containing the *gcrpr* gene in

434 other teleosts was intriguing and the presence in zebrafish of the flanking genes, *wdr90* and

435 *nrlc3*, indicate the receptor gene was potentially deleted from its genome (Figure 6). In the

436 elephant shark that lacked *gcrpr* a similar genome region within scaffold_30 was identified but

437 in this species, it flanked the putative *glp1r-like* gene.

438

439 *Gip and gcrp genome regions*

440   Comparison of the *gip* and *gcrp* gene environment in fish and humans revealed that

441 they were located close to the *hoxb* and *hoxc* cluster, respectively (Figure 7). In tilapia *gip*

442 mapped to chromosome GL831200 (2.08 Mb) and was flanked by *calcoco2* and *ttll6*. In the

443 sea bass, a similar gene complement existed and both *calcoco2* and *ttll6* were localized in close

444 proximity. However, the *gip* gene was missing suggesting it was deleted from the genome. In

445 the zebrafish genome which also contained the *gip* gene on chromosome 12 a similar gene

446 environment to that flanking tilapia *gip* was found. The *gip* neighbouring genes in the spotted

447 gar (LG15), elephant shark (sc_122) and human (chromosome 17) were similar to the tilapia.

448 However, in the spotted gar the *gip* gene was absent (Figure 7). In tilapia, the *gcrp* gene mapped

449 to GL831196 and in sea bass to LG22-25 (13,31 Mb) and homologue genome regions were

450 found in medaka chromosome 7, spotted gar LG4, coelacanth JH126563. In fish, the putative

451 *gcrp* gene was flanked by *fignl2* and *scn8* and conserved gene positions existed across

452 vertebrates (Figure 7). In human and zebrafish, the *gcrp* gene was absent but a homologue

453 genome region was found in chromosome 12 and chromosome 23, respectively.

454

455 *Tissue expression of glucagon-like members in sea bass*

The relative abundance of the glucagon-like receptors and peptide precursors was characterised in sea bass (Figure 8). Although a putative *gcrp* exon was identified it was not possible to amplify the corresponding transcript. This may be due to the small length of the sequence retrieved which limited primer design. q-PCR analysis revealed that glucagon-like receptors were low abundance and had a widespread tissue distribution. Higher transcript levels were found for *gipr*, *gcrpr* and *gcgrb* in the brain, *gcgra* in the liver and *glp2r* in the duodenum. While *gipr* and *gcrpr* expression was mainly limited to brain, *gipr* was also detected in the intestine; *gcgra*, *gcgrb* and *glp2r* were amplified in all tissues analysed (Figure 8). The gene for *proglucagon a* was mostly expressed in the intestine whereas *proglucagon b* was mostly amplified in adipose tissue (potentially containing pancreatic cells).

To assess putative function, the receptors and ligands were analysed in conditions leading to a different metabolic state. No significant changes in receptor abundance were observed in sea bass with modified plasma glucose levels (high or low) (Figure 9). The transcript abundance of the *proglucagon* precursors was significantly different in sea bass with high or low plasma glucose concentrations. *Proglucagon a* had a significantly lower expression (p < 0.05) in the intestine of sea bass with low plasma glucose concentrations (Figure 9).

4. Discussion

Members of the glucagon-like receptors and peptide precursors were identified in fish. Our results suggest that the fish glucagon-like members are related to their tetrapod homologues but have undergone a different evolutionary trajectory as reflected by their distinct gene complements and lineage specific modifications (Figure 10). The expression pattern of the receptors and precursors and their tissue abundance in sea bass suggests that they have acquired different functions. When energy metabolism was modified and glucose plasma levels were low, *proglucagon a* was down-regulated in the duodenum indicating that the link with metabolism and glucose balance has been conserved across vertebrates.

The glucagon-like system in lamprey

Lampreys are the oldest living vertebrates in which homologues of the vertebrate glucagon-like system have been found. Lampreys diverged ~500 million years ago and two *proglucagon* precursors and four *glucagon-like* receptors genes were identified two of which were duplicates of non-mammalian *gcrpr* gene. This suggests that the vertebrate receptor genes emerged early presumably during the two tetraploidization events and subsequently duplicated

489 in the radiation of the lampreys and before the sea lamprey and Japanese lamprey divergence

490 (30–10 Mya (Kuraku and Kuratani, 2006). In lampreys, despite the existence of a *glp2r* which

491 was present in all fish genomes and was highly conserved, its genome also accommodates the

492 closest fish orthologue of human *GLP1R*. Our analysis failed to identify homologues of the

493 vertebrate *gcgr* and *gipr* receptors or *gip* and *gcrp* peptide genes in the lamprey genome.

494 Previously a putative *gcgr* orthologue was suggested to exist in the sea lamprey (Irwin, 2014)

495 but our phylogenetic analysis including data from another cyclostome, the Japanese lamprey,

496 revealed that the putative lamprey *gcgr* was a paralogue of *gcrpr*. Similarly, an orthologue of

497 the vertebrate *gipr* in cyclostomes was also predicated (Irwin, 2014) but this was not retrieved

498 from either of the lamprey genome assemblies. The status of the genome assembly (short

499 genomic contigs with low gene content) means that the glucagon-like gene complement in this

500 lineage still remains to be completely defined.

501

Evolution of the glucagon-like system in other fish

503 The glucagon-like receptors and peptides in gnathostome fish are encoded by a complex

504 gene family. In the teleosts, 5 different receptors and 4 peptide precursors that can generate 7

505 different mature peptides were identified (Figure 10). The *gcgr* copy number doubled due to

506 the teleost-specific (3R) genome duplication and some species-specific gene losses such as

507 with *gipr* and *gcrpr* genes occurred suggesting that in the fishes the glucagon-like system

508 evolved into a complex and diverse family of genes in a relatively short time span. Members

509 of this family suffered a further duplication in the ancestor of the salmonids (salmonid-specific

510 autotetraploidization event (Macqueen and Johnston, 2014), but only duplicate glucagon-like

511 peptide precursors have persisted (Figure 10). In gnathostome fishes, distinct gene

512 organizations suggest that the *proglucagon* precursors evolved differently. In some fish

513 lineages, in common with lamprey the *proglucagon* gene duplicated and exon loss and

514 therefore peptide loss occurred after their divergence. In the elephant shark, exon duplication

515 occurred generating an extra exon encoding a further Glp-like peptide (that shares higher

516 sequence similarity for the vertebrate Glp2). In the coelacanth, the exon encoding the glp1

517 peptide appears to have been lost. The diversity of proglucagon precursors further supports the

518 notion that across the fishes, members of the glucagon-like system evolved differently and gene

519 persistence or gene absence was affected by distinct pressures. We propose that these include

520 the increasing complexity of the gastro-entero-pancreatic system (Youson and Al-Mahrouki,

521 1999) and the distinctive physiology of energy metabolism and homeostasis of ectotherms.

522

523   The global evolution of the fish receptors seems to mirror their putative peptide ligands

524   (Cardoso et al., 2010; Hwang et al., 2014; Irwin, 2014) with the exception of the Glp1 system,

525   as homologues of human GLP1 exist in fish but *GLP1R* is absent in some lineages. The

526   evolution of *glp1r* in fish was complex and the gene was retained in the lamprey (KE993793.1

527   in Japanese lamprey) but was subsequently lost from ray-fined fishes (spotted gar and teleost).

528   In the elephant shark the existence of this receptor was unclear and gene synteny analysis failed

529   to establish homologies for the human *GLP1R* genome region (Figure 6 and Supplementary

530   Figure 4). In the coelacanth, 5 receptors existed, four of which are orthologues of the human

531   GLP2R, GCGR, GIP and non-mammalian *gcrpr* and it is tempting to speculate that the

532   remaining receptor gene may represent a rapidly evolving Sarcopterygii *glp1r*. However,

533   localization of the receptor gene next to *gcgr* in the coelacanth genome suggests that it most

534   likely represents a putative *gcgr* gene paralogue that resulted from a tandem duplication follow

535   by gene inversion and rapid mutation (Supplementary Figure 4). We hypothesise that the

536   evolution of the *glp1r* gene in gnathostome fish was relaxed and that the receptor gene sequence

537   freely mutated and the gene was randomly translocated in the genomes which consequently

538   may have led to gene elimination in ray-finned fish genomes. Identification of orthologues of

539   the human *GLP1R* in other fish species coupled to peptide-receptor functional studies will help

540   to clarify receptor identity and evolution.

541   Curiously, the sequence homologue of the human GLP1 peptide persisted in most fish

542   genomes and activates the zebrafish Gcgrb (Oren et al., 2016). In teleosts, the paralogue Gcgra,

543   was only activated by Gcg and not by fish or human GLP1 (Chow et al., 2004) suggesting that

544   after receptor duplication functional divergence occurred and that Gcgrb acquired a dual

545   Glp1r/Gcgr function in teleost (Oren et al., 2016). The coelacanth was the only fish in which

546   no Glp1 peptide was predicted within the identified proglucagon precursor. However, it was

547   not clear if the absence of the Glp1 peptide was due to a genome misassembly or a true *glp1*

548   exon loss. If the latter is the case the possibility of an alternative peptide occupying the receptor

549   cannot be discarded.

550

551   **Unequal persistence of *gip* and *gcrp* system across fish**

552   The persistence of *gipr* and *gcrpr* and their peptide precursors in fish genomes is

553   intriguing and raises questions about their evolution and function. The reason why teleost have

554   retained a different complement of genes is unclear as the functional role of all members in fish

555 and other vertebrates is unknown. Nutritional condition is the most important physiological

556 factor regulating the glucagon-like members. Fish possess similar nutrient requirements to

557 terrestrial vertebrates but they show a wide variety of feeding habits and feeding patterns with

558 unique molecular pathways that regulate nutrition, digestion and energy stores (Volkoff and

559 Peter, 2006). In our study, no clear association between feeding behaviour and persistence and

560 loss of the *gip* or *gcrp* systems in fish genomes was identified since in both species examined,

561 the sea bass (a carnivore) and the tilapia (an herbivore/omnivore) both receptor genes were

562 retained. Salinity adaptations in fish modify glucose levels in the plasma and liver (Fiess et al.,

563 2007; Baltzegar et al., 2014) and the endocrine mechanisms that regulate energy mobilization

564 are still unclear. The species of fish used for comparative genomics ranged from highly

565 euryhaline (sea bass and stickleback), euryhaline (medaka, tilapia and Tetraodon) to

566 stenohaline (cod, zebrafish and Takifugu) but no link could be established between persistence

567 of the glucagon-like members in fish genomes with water preference.

568   With the exception of sea bass where *gipr* was present but the *gip* gene was absent, in

569 all the other teleost genomes analysed a receptor and peptide gene was identified. While

570 syntenic genome regions were found for both *gipr* and *gip* genes, the gene order of flanking

571 genes across fish was distinct suggesting that lineage specific pressures occurred. The fact that

572 the zebrafish *gipr* gene environment was more similar to that of the spotted gar and human than

573 with other teleost suggests that during their radiation gene rearrangement/shuffling occurred.

574 In contrast, gene synteny and gene order was preserved for *gcrpr* and *gcrp* across vertebrates.

575 The physiological consequence of the absence or persistence of glucagon-like family members

576 in fish genomes requires further studies. The partial characterization of a *gcrp* gene in fish

577 genomes during this study but the failure to amplify the *gcrp* precursor from sea bass raises

578 questions about its biological role. Furthermore, although *in vitro* studies indicate *gcrp* can

579 activate the *gcrpr* receptor (Park et al., 2013) its physiological role remains unclear.

580

581 **Tissue specific expression of the glucagon-like members in sea bass**

582   Expression of *gcgra* in the sea bass liver is consistent with the known primary function

583 of glucagon in regulating hepatic glycogenolysis and gluconeogenesis (Chow et al., 2004).

584 Expression of the *gcgra* paralogue (*gcgrb*) mostly in brain suggests that after gene duplication

585 functional divergence occurred and *gcg* acquired direct actions on the brain. The presence of

586 high levels of *glp2r* in the duodenum is consistent with the role of *glp2* peptide in the vertebrate

587 gastrointestinal tract (Irwin and Wong, 1995; Drucker, 2001, 2002) and detection of *gcrpr* in

588   the sea bass brain mirrors the expression of the receptor orthologues in chicken and *Xenopus*

589   (Irwin and Prentice, 2011). Expression of *gipr* was limited to brain and intestine and although

590   it was not possible to establish expression of sea bass *gip*, peptide transcripts were found in the

591   intestine and pancreas of zebrafish (Musson et al., 2009). Expression of the sea bass duplicate

592   *proglucagon* precursors also matched previous reports in fish and *proglucagon a* was mostly

593   found in the intestine while *proglucagon b* was highly expressed in the visceral adipose tissue,

594   which most likely also contained pancreatic tissue (Plisetskaya and Mommsen, 1996; Zhou and

595   Irwin, 2004). The gene expression analysis in the sea bass suggests that the intestine and the

596   visceral adipose tissue are the main organs that secrete the Gcg, Glp1 and Glp2 peptides that

597   then activate their cognate receptors in the brain, liver and intestine (Figure 11). Comparison

598   of the teleost and human glucagon-like systems revealed that the pancreatic-hepatic-intestinal

599   regulatory loop is most conserved and that the liver is the main target for Gcg and the intestine

600   is the main target of Glp2 (Figure 11). Divergence comes at the level of the intestine where a

601   putative Gcg peptide is produced in the euteleost, sea bass but not in human. The presence of

602   Gcg in fish intestine may be a remnant persisting from the more ancient fishes, which had a

603   rudimentary pancreas so that the intestine played a more important endocrine role (Youson and

604   Al-Mahrouki, 1999). The brain-gut route has been less conserved between fishes and human

605   and in humans GIP and GLPs have a primary role (Baggio and Drucker, 2007; Seino et al.,

606   2010). In fish, there is a brain-gut and adipose/pancreas–brain loop and in addition to Gip and

607   Glp, Gcg from the intestine (Gcga) and pancreas (Gcgb) may also regulate the brain through

608   Gcgrb (Figure 11). Although the affinity of the peptides for the cognate receptors of the sea

609   bass glucagon-like system was not assessed, functional characterization of the duplicate *gcgr*

610   revealed in zebrafish that the two Gcg peptides (Gcga and Gcgb) activate the paralogue Gcg

611   receptors with similar affinities (Li et al., 2015). The absence of fish *gcgr* homologue in human

612   brain and Gcg peptide from human intestine and expression of *gcrpr* in fish brain suggests that

613   this are unique characteristic of the glucagon-like system in fish. In addition, no receptor

614   expression was detected in the sea bass visceral adipose tissue where in human both GLP1R

615   and GIPR are present. Further studies should be directed at establishing ligand – receptor

616   interactions and the function of this system in fish.

617

618   **Regulatory role of the glucagon-like system**

619   Unlike mammals, the mechanism by which fish regulate glucose plasma levels is still

620   not very well understood. Across fish tolerance to glucose is variable and is dependent on food

source and environment (Moon, 2001; Navarro et al., 2002). Although carbohydrate metabolism seems to play a minor role compared to lipids and proteins, glucose plays a key role in the maintenance of metabolic homeostasis and as the main energy substrate used by several fish tissues such as brain (reviewed in (Polakof et al., 2011b). The European sea bass is a carnivore with high levels of plasma glucose. Across an annual cycle the average glucose levels are 150 mg/100 ml (8.33 mmol/L) and limiting food conditions and food depletion provoke a sharp decrease in plasma glucose levels (Gutierrez et al., 1987; Echevarria et al., 1997; Perez-Jimenez et al., 2007). Comparison of the plasma glucose levels of the experimental animals used in this study revealed that the glucose levels in the high glucose group were normal for this species (Gutierrez et al., 1987; Perez-Jimenez et al., 2007). Low levels of plasma glucose in European sea bass were associated with a significant down-regulation of *proglucagon a* in the duodenum, although no changes in the expression of the glucagon-like receptors were observed. In fish, data on *proglucagon* mRNA expression in response to feeding status is scarce however in the cyprinid Ya-fish (*Schizothorax prenanti*), the *proglucagon* precursor was also significantly down-regulated in the intestine after three days of food depletion and changes in *proglucagon* expression have been associated with regulation of food intake (Lin et al., 2015). In fish, fasting decreases Gcg and Glp1 levels in plasma, however in some species, including sea bass an increase of Gcg plasma levels during short-term fasting has been described 4 days after food depletion and then subsequently decreased (Gutierrez et al., 1991; Navarro and Gutierrez, 1995; Navarro et al., 2002). This transitory increase of the peptide in plasma has been suggested to mediate the shift from carbohydrate metabolism to the mobilization of energy stores (Gutierrez et al., 1991; Navarro et al., 1992).

In chicken, multiple *proglucagon* transcripts with distinct peptide coding potentials with a significant role in physiology have been described (Yue and Irwin, 2005; Richards and McMurtry, 2008). It remains to be established if this is also the case in fish. The role of alternative splicing in regulating peptide expression in fish is uncertain although in the rockfish (*Sebastes caurinus*), a truncated *proglucagon a* isoform that only encodes Gcg and Glp 1 was isolated from the intestine and pancreas (Busby and Mommsen, 2016).

Final considerations

The existence of different receptors and peptide precursors of the glucagon-like system in basal vertebrates' such as lamprey and elephant shark confirmed previous studies that gene members emerged and expanded early during the vertebrate radiation with the evolution of the

654 gastro-entero-pancreatic system and brain-gut regulatory loop. In lamprey, lineage-specific

655 gene duplications occurred as well as in other fish and the number of glucagon-like receptor

656 genes and genes for the ligands were variable. Putative *glp1r* genes were found in fish and the

657 *glp1r* in lamprey shared the greatest sequence similarity to the tetrapod orthologue. In teleosts

658 and spotted gar, *glp1r* was absent suggesting that the gene was eliminated early from the ray-

659 finned fish radiation and if it exists in elephant shark or coelacanth remains unresolved.

660 Homologues of the mammalian *GIP*-system and non-mammalian *gcrp*-system were found but

661 they were retained differently across fish. Unique *proglucagon* precursors were also found and

662 identification of three putative Glp-like peptides in elephant shark suggested that an extra exon

663 duplication occurred in this species. Expression of the glucagon-like system in a teleost, the

664 sea bass, revealed that they are expressed in tissues involved in metabolism and energy balance

665 and their tissue distribution suggests that they have acquired distinct functions. Changes in

666 plasma glucose levels caused by short term fasting modified the expression of one of the

667 duplicate *proglucagons* in sea bass indicating that the role of glucagon-like family members in

668 glucose homeostasis has been maintained during evolution. Considerable work will be required

669 to clearly establish the role in fish of the expanded glucagon-like system repertoire identified

670 in the brain-gut regulatory axis.

671

677

678

679

680

681 **Tables:**

682 **Table 1:** Glucagon and glucagon-like 1 peptides and *proglucagon* cDNA isolated from fish.

683 The length (aa) of the isolated peptides and *proglucagon* isoforms (*a* and *b* or I and II in the

684 case of lamprey) and tissue of origin are indicated.

685

686 **Table 2:** Primer sequences used to amplify the glucagon-like receptors and ligands in sea bass.

688 **Figure legends:**

689 **Figure 1:** The fish glucagon-like receptors and peptide genes. The cladogram includes the
690 number of genes retrieved from each of the fish genomes analysed confirmed by phylogeny.
691 The teleost genes are highlighted by a grey box. Human was included for comparative
692 purposes. The teleost specific genome duplication (TSGD) is indicated by a black dot and genes
693 that were not identified are indicated by "ni". Tilapia and sea bass are the only teleost where 5
694 glucagon-like receptor genes were identified and they are boxed. * very incomplete sequence;
695 ? unclear if present.

696

697 **Figure 2:** Alignment of the elephant shark, spotted gar, coelacanth and human deduced
698 proglucagons. The deduced peptides are annotated by coloured boxes and peptide sequence is
699 in italics. The proteolytic cleavage sites (mono basic or dibasic) that will generate the predicted
700 peptides are in bold. The additional Glp-like encoded in the elephant shark precursor is
701 underlined. * (asterisk) indicates fully conserved residues. Dashes (–) are gaps inserted to
702 maximize the alignment. The consensus sequence was manually edited to evidence the
703 conservation between the fish and human Glp1 peptides. The human signal peptide is doubled
704 underlined.

705

706 **Figure 3:** Organization of the fish and human *proglucagon* gene. Only the coding exons are
707 represented and the exons encoding the glucagon-like peptides are represented by different
708 shapes and colours (Gcg, green; Glp1, pink; Glp2 blue). The predicted sizes (aa) of the mature
709 peptides are indicated below the exons. The signal peptide exon (sp, coloured in red) and other
710 non-mature peptide coding exons (coloured in white) are indicated. Solid lines represent
711 introns. The figure is not drawn to scale.

712

713 **Figure 4:** Phylogenetic tree of the fish glucagon-like receptors. The tree was obtained using
714 the BI method and the posterior probability values of the main vertebrate clades are indicated.
715 The five main receptor clades are annotated and tilapia and sea bass receptors are in bold. The
716 tree was constructed using the deduced receptor sequences from TM1 to TM7. The tree was
717 rooted with human PTHR1. A similar tree was constructed with the ML method and 100
718 bootstrap and is available as Supplementary Figure 1.

719

Figure 5: Phylogenetic tree of the fish glucagon-like peptides. The tree was obtained using the BI method and posterior probability values of the main clades are indicated. The cladogram tree is available as Supplementary Figure 2. The five main peptide clades are annotated and the tilapia and sea bass peptides are highlighted in bold. The tree was constructed with the conserved mature peptide (1-28) sequence alignment. A similar tree constructed using the ML method with 100 bootstrap was obtained and is available in Supplementary File 3. The sequence of the predicted medaka *gcrp* was very incomplete and was not used. The putative lamprey Gip described by (Musson et al., 2011) was also omitted as no corresponding gene was found in the current sea lamprey genome assembly. Salmon duplicate proglucagon peptides were designated Gcg aa and Gcg ab as they tend to group within the teleost Gcg a cluster and the duplicate Gip were named Gip a and Gip b, as this is the first teleost where duplicates of this peptide were found.

Figure 6: Gene environment comparison of *gipr* and *gcrpr* in fish and human. The neighbouring gene environment of the sea bass receptors was characterised (grey background) and used to identify the homologue genome regions in other vertebrates. Block arrows correspond to genes (HUGO annotation) and the arrowheads point in the direction of gene transcription. Dashed arrows represent genes that are not annotated in the databases but were found by similarity searches. Gene homologues in different species are denoted by the same colour and they are aligned. Horizontal lines represent chromosome fragments and the relative position of genes is given (Mb). Numbers indicate members of the same family in the same genome fragment. Dashed box group the genes that were potentially inserted in the neighbourhood of the teleost *gipr* as they are localized on non-homologous genome/chromosome regions in other vertebrates.

Figure 7: Comparison of the gene environment of the fish and human *gip* and *gcrp* genome regions. The neighbouring gene environment of the tilapia *gip* and sea bass *gcrp* (both with a grey background) were characterized and used to identify homologue genome regions in other vertebrates. Block arrows correspond to genes (HUGO annotation) and arrowheads point in the direction of gene transcription. Gene homologues are denoted by the same colour and they are aligned. Horizontal lines represent chromosome fragments and the relative gene positions is given (Mb). The *hox* gene clusters are represented by lined-filled squared. The elephant shark

752   glucagon-like gene found in the homologous genome region is represented by a dashed box

753   and corresponds to a *glp1r-like* gene based on its higher similarity for the human GLP1R.

754

755   **Figure 8:** Tissue distribution of transcripts of the glucagon-like receptors and proglucagon in

756   sea bass. Receptor expression levels were obtained by q-PCR and was normalized using the

757   geometric mean of two reference genes (Elongation factor 1-alpha, *ef1α* and 18S Ribossomal

758   RNA, *18s*). The data corresponds to the mean ± SEM (n = 3, biological replicates).

759

760   **Figure 9:** Expression of the glucagon-like receptors and proglucagon in target tissues of food

761   challenged sea bass. A) Plasma glucose levels (mmol/L) (n = 6 biological replicates, ** p

762   <0.0001); B) Variation of transcript expression was determined by q-PCR and normalized

763   using the geometric mean of two reference genes (Elongation factor 1-alpha (*ef1α*) and 18S

764   Ribossomal RNA (*18s*)). Target tissues were selected based on transcript abundance (Figure

765   8). Data is presented as the mean ± SEM (n = 6 biological replicates) and statistical significance

766   was considered at p < 0.05 (*, two-tailed unpaired Student's t-test).

767

768   **Figure 10:** Proposed evolutionary model for the glucagon-like system in fish. Genes for

769   peptides and receptors are proposed to have emerged by gene and exon duplication events early

770   at the emergence of vertebrates. The two genome duplication events at the origin of vertebrates

771   (1R and 2R) and the teleost specific genome duplication event (3R) are indicated. Receptors

772   and pro-peptides are coloured. The teleost clade is highlighted in grey. Only the genes that

773   were identified are represented and "X" indicates potential gene deletion during the vertebrate

774   evolution. Duplicate genes are denoted by a and b or by I and II in the case of the lamprey and

775   "?" represents unclear existence. In salmon, the duplicate *proglucagon* (*aa* and *ab*) and *gip* (*a*

776   and *b*) are likely to be the result of the specific salmonid tetraploidization (not represented). B)

777   Diagrammatic representation of the distribution of the entero-hepato-pancreas system in fish

778   (adult lamprey, holocephalian, basal actinopterygian and eutelost, adapted from (Youson and

779   Al-Mahrouki, 1999) is included to exemplify its increasing complexity from extant agnatha to

780   advanced teleost. In lamprey, the intestine contains both exocrine and endocrine cells. With the

781   radiation of the fish an exocrine pancreas emerged. For the euteleosts the images represent both

782   gastric (with a functional stomach stickleback, sea bass, tilapia, Atlantic cod and Atlantic

783   salmon) and agastric (zebrafish, medaka) species: G- gall bladder, L- liver, S- stomach

784

785   **Figure 11:** Distribution and potential interactions of the members of the glucagon-like system
786   in fish and human. Localization of glucagon receptors and peptides were mapped to brain, liver,
787   intestine and pancreas/adipose, organs that are involved in the regulation of vertebrate
788   metabolism and energy homeostasis. Fish expression data was obtained from Figure 8 and
789   transcripts were mapped to the tissues where they were found most abundant and also from
790   Table 1 and for the teleost *gip* from (Musson et al., 2011). Human data was obtained from
791   Human protein atlas (http://www.proteinatlas.org) and from (Gremlich et al., 1995; Wei and
792   Mojsov, 1995). Receptors and their likely activating peptides are shaded in the same colour
793   with the exception of *proglucagon* (*a* and *b*) transcripts in the fish. The arrows point to potential
794   sites of action of the mature peptides as the receptor-peptide interactions of the duplicated fish
795   system still remains poorly characterized. The fish and human systems seem to overlap and
796   pancreas/visceral adipose and intestine are the main tissues where glucagon-like peptide
797   members are secreted and mainly act on liver, brain and intestine. The human pancreas mostly
798   secretes GCG while in fish both Gcg and Glp1 peptides (derived from *proglucagon b*) seem to
799   exist (Table 1). Expression of the *gcrp* precursor in fish remains to be established. The *gcgrb*
800   expressed in the brain may have a dual Gcgr/Glp1r role in fish as described for the zebrafish
801   homologue (Oren et al., 2016).
802
803   **Supplementary data**
804   **Supplementary Figure 1:** ML phylogenetic tree of the glucagon-like receptors. Tree was
805   performed according to the parameters described in the methods. Bootstrap values lower than
806   50 (< 50) were deleted. Tilapia and sea bass receptors are highlighted in bold.
807
808   **Supplementary Figure 2:** BI phylogenetic tree (cladogram) of the glucagon-like peptides.
809   Tree was performed according to the parameters described in the methods.
810
811   **Supplementary Figure 3:** ML phylogenetic tree of the glucagon-like peptides. Tree was
812   performed according to the parameters described in the methods
813
814   **Supplementary Figure 4:** Gene synteny of the coelacanth glucagon-like receptor genome
815   region with the human. The human *GLP1R* gene environment on chromosome 6 is also
816   represented and was mapped on fish. The spotted gar was used as the ray-finned fish
817   representative which lost *glp1r* gene. The putative elephant shark *glp1r* maps to scaffold 30

818   that possess a conserved gene environment with the fish *gcrpr* genome region (Figure 6). No

819   data was obtained from lamprey genome where a putative *glp1r* was confirmed by phylogeny.

820   The coelacanth glucagon-like member is dashed has its identity remains to be further clarified.

821

822   **Supplementary File 1:** Glucagon-like receptor sequences used for phylogenetic analysis.

823   Sequence from TM1 to TM7 including intracellular and extracellular loops were used. The

824   sequence fragments in bold were miss aligned and were removed from the edited alignment.

825

826   **Supplementary File 2:** Amino acid sequences of the glucagon-like peptide precursors. The

827   putative proteolytic cleavage sites that generate the deduced mature peptides are annotated in

828   bold and the mature peptides used for phylogeny is in italic. Underlined are the elephant shark

829   Glp-like peptides. The coelacanth proglucagon does not encode for Glp1. The platyfish

830   proglucagon a precursor is incomplete and lacks Glp2. The lizard proglucagon is probably

831   incomplete and encodes only for Gcg. Lamprey proglucagon I encode for Gcg, Glp1 and Glp2

832   and proglucagon II Gcg and Glp2.

833

834   **Supplementary File 3:** Amino acid sequences used for the peptide phylogenetic analysis. A

835   a and b indicate the peptides derived from proglucagon a and proglucagon b, respectively.

836   Similar for the lamprey proglucagon I and II.

837

838   **Supplementary Table 1:** Accession numbers of the chordate glucagon-like receptors.

839   Searches for the fish receptors were mostly performed in ENSEMBL with some exceptions

840   (see text). When no gene was predicted the NCBI accession number when available is given

841   and also its genome localization; n.i.-not identified

842

843   **Supplementary Table 2:** Accession numbers of the glucagon-like peptide precursors.

844   Searches for the fish genes were performed in ENSEMBL with some exceptions (see text).

845   When no gene was predicted the gene localization is given. * EST; n.i.-not identified

846

847

848

849

850

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104-2105.

Albalat, A., Gomez-Requeni, P., Rojas, P., Medale, F., Kaushik, S., Vianen, G.J., Van den Thillart, G., Gutierrez, J., Perez-Sanchez, J., Navarro, I., 2005. Nutritional and hormonal control of lipolysis in isolated gilthead seabream (Sparus aurata) adipocytes. Am J Physiol Regul Integr Comp Physiol 289, R259-265.

Amemiya, C.T., Alfoldi, J., Lee, A.P., Fan, S.H., Philippe, H., MacCallum, I., Braasch, I., Manousaki, T., Schneider, I., Rohner, N., Organ, C., Chalopin, D., Smith, J.J., Robinson, M., Dorrington, R.A., Gerdol, M., Aken, B., Biscotti, M.A., Barucca, M., Baurain, D., Berlin, A.M., Blatch, G.L., Buonocore, F., Burmester, T., Campbell, M.S., Canapa, A., Cannon, J.P., Christoffels, A., De Moro, G., Edkins, A.L., Fan, L., Fausto, A.M., Feiner, N., Forconi, M., Gamieldien, J., Gnerre, S., Gnirke, A., Goldstone, J.V., Haerty, W., Hahn, M.E., Hesse, U., Hoffmann, S., Johnson, J., Karchner, S.I., Kuraku, S., Lara, M., Levin, J.Z., Litman, G.W., Mauceli, E., Miyake, T., Mueller, M.G., Nelson, D.R., Nitsche, A., Olmo, E., Ota, T., Pallavicini, A., Panji, S., Picone, B., Ponting, C.P., Prohaska, S.J., Przybylski, D., Saha, N.R., Ravi, V., Ribeiro, F.J., Sauka-Spengler, T., Scapigliati, G., Searle, S.M.J., Sharpe, T., Simakov, O., Stadler, P.F., Stegeman, J.J., Sumiyama, K., Tabbaa, D., Tafer, H., Turner-Maier, J., van Heusden, P., White, S., Williams, L., Yandell, M., Brinkmann, H., Volff, J.N., Tabin, C.J., Shubin, N., Schartl, M., Jaffe, D.B., Postlethwait, J.H., Venkatesh, B., Di Palma, F., Lander, E.S., Meyer, A., Lindblad-Toh, K., 2013. The African coelacanth genome provides insights into tetrapod evolution. Nature 496, 311-316.

Baggio, L.L., Drucker, D.J., 2007. Biology of incretins: GLP-1 and GIP. Gastroenterology 132, 2131-2157.

Baltzegar, D.A., Reading, B.J., Douros, J.D., Borski, R.J., 2014. Role for leptin in promoting glucose mobilization during acute hyperosmotic stress in teleost fishes. J Endocrinol 220, 61-72.

Bell, G.I., 1986. The glucagon superfamily: precursor structure and gene organization. Peptides 7 Suppl 1, 27-36.

Bortolato, A., Dore, A.S., Hollenstein, K., Tehan, B.G., Mason, J.S., Marshall, F.H., 2014. Structure of Class B GPCRs: new horizons for drug discovery. Br J Pharmacol 171, 3132-3145.

Braasch, I., Gehrke, A.R., Smith, J.J., Kawasaki, K., Manousaki, T., Pasquier, J., Amores, A., Desvignes, T., Batzel, P., Catchen, J., Berlin, A.M., Campbell, M.S., Barrell, D., Martin, K.J., Mulley, J.F., Ravi, V., Lee, A.P., Nakamura, T., Chalopin, D., Fan, S., Wcisel, D., Canestro, C., Sydes, J., Beaudry, F.E., Sun, Y., Hertel, J., Beam, M.J., Fasold, M., Ishiyama, M., Johnson, J., Kehr, S., Lara, M., Letaw, J.H., Litman, G.W., Litman, R.T., Mikami, M., Ota, T., Saha, N.R., Williams, L., Stadler, P.F., Wang, H., Taylor, J.S., Fontenot, Q., Ferrara, A., Searle, S.M., Aken, B., Yandell, M., Schneider, I., Yoder, J.A., Volff, J.N., Meyer, A., Amemiya, C.T., Venkatesh, B., Holland, P.W., Guiguen, Y., Bobe, J., Shubin, N.H., Di Palma, F., Alfoldi, J., Lindblad-Toh, K., Postlethwait, J.H., 2016. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. Nat Genet 48, 427-437.

Buchan, A.M., Polak, J.M., Capella, C., Solcia, E., Pearse, A.G., 1978. Electronimmunocytochemical evidence for the K cell localization of gastric inhibitory polypeptide (GIP) in man. Histochemistry 56, 37-44.

Busby, E.R., Mommsen, T.P., 2016. Proglucagons in vertebrates: Expression and processing of multiple genes in a bony fish. Comp Biochem Physiol B Biochem Mol Biol 199, 58-66.

Campbell, R.M., Scanes, C.G., 1992. Evolution of the growth hormone-releasing factor (GRF) family of peptides. Growth Regul 2, 175-191.

900 Campos, R.V., Lee, Y.C., Drucker, D.J., 1994. Divergent tissue-specific and developmental
901 expression of receptors for glucagon and glucagon-like peptide-1 in the mouse. Endocrinology
902 134, 2156-2164.
903 Cardoso, J.C., Bergqvist, C.A., Felix, R.C., Larhammar, D., 2016. Corticotropin-releasing
904 hormone family evolution: five ancestral genes remain in some lineages. J Mol Endocrinol 57,
905 73-86.
906 Cardoso, J.C., Clark, M.S., Viera, F.A., Bridge, P.D., Gilles, A., Power, D.M., 2005. The
907 secretin G-protein-coupled receptor family: teleost receptors. J Mol Endocrinol 34, 753-765.
908 Cardoso, J.C., Felix, R.C., Bergqvist, C.A., Larhammar, D., 2014a. New insights into the
909 evolution of vertebrate CRH (corticotropin-releasing hormone) and invertebrate DH44
910 (diuretic hormone 44) receptors in metazoans. Gen Comp Endocrinol 209, 162-170.
911 Cardoso, J.C., Felix, R.C., Trindade, M., Power, D.M., 2014b. Fish genomes provide novel
912 insights into the evolution of vertebrate secretin receptors and their ligand. Gen Comp
913 Endocrinol 209, 82-92.
914 Cardoso, J.C., Vieira, F.A., Gomes, A.S., Power, D.M., 2010. The serendipitous origin of
915 chordate secretin peptide family members. BMC Evol Biol 10, 135.
916 Chen, S., Zhang, G., Shao, C., Huang, Q., Liu, G., Zhang, P., Song, W., An, N., Chalopin, D.,
917 Volff, J.N., Hong, Y., Li, Q., Sha, Z., Zhou, H., Xie, M., Yu, Q., Liu, Y., Xiang, H., Wang, N.,
918 Wu, K., Yang, C., Zhou, Q., Liao, X., Yang, L., Hu, Q., Zhang, J., Meng, L., Jin, L., Tian, Y.,
919 Lian, J., Yang, J., Miao, G., Liu, S., Liang, Z., Yan, F., Li, Y., Sun, B., Zhang, H., Zhang, J.,
920 Zhu, Y., Du, M., Zhao, Y., Schartl, M., Tang, Q., Wang, J., 2014. Whole-genome sequence of
921 a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic
922 lifestyle. Nat Genet 46, 253-260.
923 Chow, B.K., Moon, T.W., Hoo, R.L., Yeung, C.M., Muller, M., Christos, P.J., Mojsov, S.,
924 2004. Identification and characterization of a glucagon receptor from the goldfish Carassius
925 auratus: implications for the evolution of the ligand specificity of glucagon receptors in
926 vertebrates. Endocrinology 145, 3273-3288.
927 Christophe, J., 1996. Glucagon and its receptor in various tissues. Ann N Y Acad Sci 805, 31-
928 42; discussion 42-33.
929 Conlon, J.M., Andrews, P.C., Thim, L., Moon, T.W., 1991. The primary structure of glucagon-
930 like peptide but not insulin has been conserved between the American eel, Anguilla rostrata
931 and the European eel, Anguilla anguilla. Gen Comp Endocrinol 82, 23-32.
932 Conlon, J.M., Nielsen, P.F., Youson, J.H., 1993a. Primary structures of glucagon and glucagon-
933 like peptide isolated from the intestine of the parasitic phase lamprey Petromyzon marinus.
934 Gen Comp Endocrinol 91, 96-104.
935 Conlon, J.M., Youson, J.H., Mommsen, T.P., 1993b. Structure and biological activity of
936 glucagon and glucagon-like peptide from a primitive bony fish, the bowfin (Amia calva).
937 Biochem J 295 (Pt 3), 857-861.
938 D'Alessio, D.A., Fujimoto, W.Y., Ensinck, J.W., 1989. Effects of glucagonlike peptide I-(7-
939 36) on release of insulin, glucagon, and somatostatin by rat pancreatic islet cell monolayer
940 cultures. Diabetes 38, 1534-1538.
941 Davidson, W.S., Koop, B.F., Jones, S.J., Iturra, P., Vidal, R., Maass, A., Jonassen, I., Lien, S.,
942 Omholt, S.W., 2010. Sequencing the genome of the Atlantic salmon (Salmo salar). Genome
943 Biol 11, 403.
944 Donahey, J.C., van Dijk, G., Woods, S.C., Seeley, R.J., 1998. Intraventricular GLP-1 reduces
945 short- but not long-term food intake or body weight in lean and obese rats. Brain Res 779, 75-
946 83.
947 Drucker, D.J., 2001. Glucagon-like peptide 2. J Clin Endocrinol Metab 86, 1759-1764.
948 Drucker, D.J., 2002. Gut adaptation and the glucagon-like peptides. Gut 50, 428-435.

949 Drucker, D.J., Erlich, P., Asa, S.L., Brubaker, P.L., 1996. Induction of intestinal epithelial
950 proliferation by glucagon-like peptide 2. Proc Natl Acad Sci U S A 93, 7911-7916.

951 Drucker, D.J., Philippe, J., Mojsov, S., Chick, W.L., Habener, J.F., 1987. Glucagon-like
952 peptide I stimulates insulin gene expression and increases cyclic AMP levels in a rat islet cell
953 line. Proc Natl Acad Sci U S A 84, 3434-3438.

954 Echevarria, G., MartinezBebia, M., Zamora, S., 1997. Evolution of biometric indices and
955 plasma metabolites during prolonged starvation in European sea bass (Dicentrarchus labrax,
956 L). Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology 118,
957 111-123.

958 Eckel, R.H., Fujimoto, W.Y., Brunzell, J.D., 1979. Gastric inhibitory polypeptide enhanced
959 lipoprotein lipase activity in cultured preadipocytes. Diabetes 28, 1141-1142.

960 Estall, J.L., Drucker, D.J., 2006. Glucagon-like Peptide-2. Annu Rev Nutr 26, 391-411.

961 Fiess, J.C., Kunkel-Patterson, A., Mathias, L., Riley, L.G., Yancey, P.H., Hirano, T., Grau,
962 E.G., 2007. Effects of environmental salinity and temperature on osmoregulatory ability,
963 organic osmolytes, and plasma hormone profiles in the Mozambique tilapia (Oreochromis
964 mossambicus). Comp Biochem Physiol A Mol Integr Physiol 146, 252-264.

965 Gremlich, S., Porret, A., Hani, E.H., Cherif, D., Vionnet, N., Froguel, P., Thorens, B., 1995.
966 Cloning, functional expression, and chromosomal localization of the human pancreatic islet
967 glucose-dependent insulinotropic polypeptide receptor. Diabetes 44, 1202-1208.

968 Guerreiro, P.M., Renfro, J.L., Power, D.M., Canario, A.V., 2007. The parathyroid hormone
969 family of peptides: structure, tissue distribution, regulation, and potential functional roles in
970 calcium and phosphate balance in fish. Am J Physiol Regul Integr Comp Physiol 292, R679-
971 696.

972 Gutierrez, J., Fernandez, J., Carrillo, M., Zanuy, S., Planas, J., 1987. Annual cycle of plasma
973 insulin and glucose of sea bass.Dicentrarchus labrax, L. Fish Physiol Biochem 4, 137-141.

974 Gutierrez, J., Perez, J., Navarro, I., Zanuy, S., Carrillo, M., 1991. Changes in plasma glucagon
975 and insulin associated with fasting in sea bass (Dicentrarchus labrax). Fish Physiol Biochem 9,
976 107-112.

977 Harmar, A.J., 2001. Family-B G-protein-coupled receptors. Genome Biol 2, REVIEWS3013.

978 Harmon, J.S., Sheridan, M.A., 1992. Effects of nutritional state, insulin, and glucagon on lipid
979 mobilization in rainbow trout, Oncorhynchus mykiss. Gen Comp Endocrinol 87, 214-221.

980 Henkel, C.V., Dirks, R.P., de Wijze, D.L., Minegishi, Y., Aoyama, J., Jansen, H.J., Turner, B.,
981 Knudsen, B., Bundgaard, M., Hvam, K.L., Boetzer, M., Pirovano, W., Weltzien, F.A., Dufour,
982 S., Tsukamoto, K., Spaink, H.P., van den Thillart, G.E., 2012. First draft genome sequence of
983 the Japanese eel, Anguilla japonica. Gene 511, 195-201.

984 Holst, J.J., 2007. The physiology of glucagon-like peptide 1. Physiol Rev 87, 1409-1439.

985 Holst, J.J., Orskov, C., Nielsen, O.V., Schwartz, T.W., 1987. Truncated glucagon-like peptide
986 I, an insulin-releasing hormone from the distal gut. FEBS Lett 211, 169-174.

987 Hoyle, C.H., 1998. Neuropeptide families: evolutionary perspectives. Regul Pept 73, 1-33.

988 Hwang, J.I., Moon, M.J., Park, S., Kim, D.K., Cho, E.B., Ha, N., Son, G.H., Kim, K., Vaudry,
989 H., Seong, J.Y., 2013. Expansion of secretin-like G protein-coupled receptors and their peptide
990 ligands via local duplications before and after two rounds of whole-genome duplication. Mol
991 Biol Evol 30, 1119-1130.

992 Hwang, J.I., Yun, S., Moon, M.J., Park, C.R., Seong, J.Y., 2014. Molecular evolution of
993 GPCRs: GLP1/GLP1 receptors. J Mol Endocrinol 52, T15-27.

994 Irwin, D.M., 2001. Molecular evolution of proglucagon. Regul Pept 98, 1-12.

995 Irwin, D.M., 2002. Ancient duplications of the human proglucagon gene. Genomics 79, 741-
996 746.

997 Irwin, D.M., 2014. Evolution of receptors for peptides similar to glucagon. Gen Comp
998 Endocrinol 209, 50-60.

999 Irwin, D.M., Huner, O., Youson, J.H., 1999. Lamprey proglucagon and the origin of glucagon-
1000 like peptides. Mol Biol Evol 16, 1548-1557.
1001 Irwin, D.M., Prentice, K.J., 2011. Incretin hormones and the expanding families of glucagon-
1002 like sequences and their receptors. Diabetes Obes Metab 13 Suppl 1, 69-81.
1003 Irwin, D.M., Wong, J., 1995. Trout and chicken proglucagon: alternative splicing generates
1004 mRNA transcripts encoding glucagon-like peptide 2. Mol Endocrinol 9, 267-277.
1005 Irwin, D.M., Zhang, T., 2006. Evolution of the vertebrate glucose-dependent insulinotropic
1006 polypeptide (GIP) gene. Comp Biochem Physiol Part D Genomics Proteomics 1, 385-395.
1007 Jiang, G., Zhang, B.B., 2003. Glucagon and regulation of glucose metabolism. Am J Physiol
1008 Endocrinol Metab 284, E671-678.
1009 Kieffer, T.J., Habener, J.F., 1999. The glucagon-like peptides. Endocr Rev 20, 876-913.
1010 Komatsu, R., Matsuyama, T., Namba, M., Watanabe, N., Itoh, H., Kono, N., Tarui, S., 1989.
1011 Glucagonostatic and insulinotropic action of glucagonlike peptide I-(7-36)-amide. Diabetes 38,
1012 902-905.
1013 Kuraku, S., Kuratani, S., 2006. Time scale for cyclostome evolution inferred with a
1014 phylogenetic diagnosis of hagfish and lamprey cDNA sequences. Zoolog Sci 23, 1053-1064.
1015 Lagerstrom, M.C., Schioth, H.B., 2008. Structural diversity of G protein-coupled receptors and
1016 significance for drug discovery. Nat Rev Drug Discov 7, 339-357.
1017 Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large
1018 datasets. Bioinformatics 30, 3276-3278.
1019 Li, M., Dean, E.D., Zhao, L., Nicholson, W.E., Powers, A.C., Chen, W., 2015. Glucagon
1020 receptor inactivation leads to alpha-cell hyperplasia in zebrafish. J Endocrinol 227, 93-103.
1021 Lin, F., Chen, H., Liu, H., Gao, Y., Zhang, X., Hao, J., Chen, D., Wu, H., Yuan, D., Wang, T.,
1022 Li, Z., 2015. Molecular cloning of a proglucagon in a cyprinid fish (Schizothorax prenanti):
1023 mRNA tissue distribution and qu
1024 antification during periprandial changes and fasting. Aquaculture 448, 250-255.
1025 Louis, A., Muffato, M., Roest Crollius, H., 2013. Genomicus: five genome browsers for
1026 comparative genomics in eukaryota. Nucleic Acids Res 41, D700-705.
1027 Lund, P.K., Goodman, R.H., Dee, P.C., Habener, J.F., 1982. Pancreatic preproglucagon cDNA
1028 contains two glucagon-related coding sequences arranged in tandem. Proc Natl Acad Sci U S
1029 A 79, 345-349.
1030 Macqueen, D.J., Johnston, I.A., 2014. A well-constrained estimate for the timing of the
1031 salmonid whole genome duplication reveals major decoupling from species diversification.
1032 Proc Biol Sci 281, 20132881.
1033 Martins, R., Vieira, F.A., Power, D.M., 2014. Calcitonin receptor family evolution and fishing
1034 for function using in silico promoter analysis. Gen Comp Endocrinol 209, 61-73.
1035 Mehta, T.K., Ravi, V., Yamasaki, S., Lee, A.P., Lian, M.M., Tay, B.H., Tohari, S., Yanai, S.,
1036 Tay, A., Brenner, S., Venkatesh, B., 2013. Evidence for at least six Hox clusters in the Japanese
1037 lamprey (Lethenteron japonicum). Proc Natl Acad Sci U S A 110, 16044-16049.
1038 Meier, J.J., Nauck, M.A., 2005. Glucagon-like peptide 1(GLP-1) in biology and pathology.
1039 Diabetes Metab Res Rev 21, 91-117.
1040 Mojsov, S., 2000. Glucagon-like peptide-1 (GLP-1) and the control of glucose metabolism in
1041 mammals and teleost fish. American Zoologist 40, 246-258.
1042 Mojsov, S., Heinrich, G., Wilson, I.B., Ravazzola, M., Orci, L., Habener, J.F., 1986.
1043 Preproglucagon gene expression in pancreas and intestine diversifies at the level of post-
1044 translational processing. J Biol Chem 261, 11880-11889.
1045 Mojsov, S., Weir, G.C., Habener, J.F., 1987. Insulinotropin: glucagon-like peptide I (7-37) co-
1046 encoded in the glucagon gene is a potent stimulator of insulin release in the perfused rat
1047 pancreas. J Clin Invest 79, 616-619.

1048   Mommsen, T.P., Andrews, P.C., Plisetskaya, E.M., 1987. Glucagon-like peptides activate
1049   hepatic gluconeogenesis. FEBS Lett 219, 227-232.
1050   Mommsen, T.P., Moon, T.W., 1989. Metabolic actions of glucagon-family hormones in liver.
1051   Fish Physiol Biochem 7, 279-288.
1052   Moon, T.W., 2001. Glucose intolerance in teleost fish: fact or fiction? Comp Biochem Physiol
1053   B Biochem Mol Biol 129, 243-249.
1054   Musson, M.C., Jepeal, L.I., Finnerty, J.R., Wolfe, M.M., 2011. Evolutionary expression of
1055   glucose-dependent-insulinotropic polypeptide (GIP). Regul Pept 171, 26-34.
1056   Musson, M.C., Jepeal, L.I., Mabray, P.D., Zhdanova, I.V., Cardoso, W.V., Wolfe, M.M., 2009.
1057   Expression of glucose-dependent insulinotropic polypeptide in the zebrafish. Am J Physiol
1058   Regul Integr Comp Physiol 297, R1803-1812.
1059   Navarro, I., Gutierrez, J., 1995. Fasting and starvation, in: Mommsen, H.a. (Ed.), Biochemistry
1060   and molecular biology of fishes. Elsevier Science B.V., pp. 393-434.
1061   Navarro, I., Gutierrez, J., Planas, J., 1992. Changes in plasma glucagon, insulin and tissue
1062   metabolites associated with prolonged fasting in brown trout (Salmo trutta fario) during two
1063   different seasons of the year. Comp Biochem Physiol Comp Physiol 102, 401-407.
1064   Navarro, I., Rojas, P., Capilla, E., Albalat, A., Castillo, J., Montserrat, N., Codina, M.,
1065   Gutierrez, J., 2002. Insights into insulin and glucagon responses in fish. Fish Physiology and
1066   Biochemistry 27, 205-216.
1067   Ng, S.Y., Lee, L.T., Chow, B.K., 2010. Insights into the evolution of proglucagon-derived
1068   peptides and receptors in fish and amphibians. Ann N Y Acad Sci 1200, 15-32.
1069   Oren, D.A., Wei, Y., Skrabanek, L., Chow, B.K., Mommsen, T., Mojsov, S., 2016. Structural
1070   Mapping and Functional Characterization of Zebrafish Class B G-Protein Coupled Receptor
1071   (GPCR) with Dual Ligand Selectivity towards GLP-1 and Glucagon. PLoS One 11, e0167718.
1072   Park, C.R., Moon, M.J., Park, S., Kim, D.K., Cho, E.B., Millar, R.P., Hwang, J.I., Seong, J.Y.,
1073   2013. A novel glucagon-related peptide (GCRP) and its receptor GCRPR account for
1074   coevolution of their family members in vertebrates. PLoS One 8, e65420.
1075   Pederson, R.A., Schubert, H.E., Brown, J.C., 1975. Gastric inhibitory polypeptide. Its
1076   physiologic release and insulinotropic action in the dog. Diabetes 24, 1050-1056.
1077   Perez-Jimenez, A., Guedes, M.J., Morales, A.E., Oliva-Teles, A., 2007. Metabolic responses
1078   to short starvation and refeeding in Dicentrarchus labrax. Effect of dietary composition.
1079   Aquaculture 265, 325-335.
1080   Plisetskaya, E.M., Mommsen, T.P., 1996. Glucagon and glucagon-like peptides in fishes. Int
1081   Rev Cytol 168, 187-257.
1082   Plisetskaya, E.M., Ottolenghi, C., Sheridan, M.A., Mommsen, T.P., Gorbman, A., 1989.
1083   Metabolic effects of salmon glucagon and glucagon-like peptide in coho and chinook salmon.
1084   Gen Comp Endocrinol 73, 205-216.
1085   Polakof, S., Miguez, J.M., Soengas, J.L., 2011a. Evidence for a Gut-Brain Axis Used by
1086   Glucagon-like Peptide-1 to Elicit Hyperglycaemia in Fish. Journal of Neuroendocrinology 23,
1087   508-518.
1088   Polakof, S., Mommsen, T.P., Soengas, J.L., 2011b. Glucosensing and glucose homeostasis:
1089   from fish to mammals. Comp Biochem Physiol B Biochem Mol Biol 160, 123-149.
1090   Pyke, C., Heller, R.S., Kirk, R.K., Orskov, C., Reedtz-Runge, S., Kaastrup, P., Hvelplund, A.,
1091   Bardram, L., Calatayud, D., Knudsen, L.B., 2014. GLP-1 receptor localization in monkey and
1092   human tissue: novel distribution revealed with extensively validated monoclonal antibody.
1093   Endocrinology 155, 1280-1290.
1094   Ramnanan, C.J., Edgerton, D.S., Kraft, G., Cherrington, A.D., 2011. Physiologic action of
1095   glucagon on liver glucose metabolism. Diabetes Obes Metab 13 Suppl 1, 118-125.
1096   Ravi, V., Venkatesh, B., 2008. Rapidly evolving fish genomes and teleost diversity. Curr Opin
1097   Genet Dev 18, 544-550.

1098   Richards, M.P., McMurtry, J.P., 2008. Expression of proglucagon and proglucagon-derived
1099   peptide hormone receptor genes in the chicken. Gen Comp Endocrinol 156, 323-338.
1100   Roch, G.J., Wu, S., Sherwood, N.M., 2009. Hormones and receptors in fish: do duplicates
1101   matter? Gen Comp Endocrinol 161, 3-12.
1102   Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B.,
1103   Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic
1104   inference and model choice across a large model space. Syst Biol 61, 539-542.
1105   Ross, S.A., Dupre, J., 1978. Effects of ingestion of triglyceride or galactose on secretion of
1106   gastric inhibitory polypeptide and on responses to intravenous glucose in normal and diabetic
1107   subjects. Diabetes 27, 327-333.
1108   Seino, Y., Fukushima, M., Yabe, D., 2010. GIP and GLP-1, the two incretin hormones:
1109   Similarities and differences. J Diabetes Investig 1, 8-23.
1110   Sherwood, N.M., Krueckl, S.L., McRory, J.E., 2000. The origin and function of the pituitary
1111   adenylate cyclase-activating polypeptide (PACAP)/glucagon superfamily. Endocr Rev 21,
1112   619-670.
1113   Silverstein, J.T., Bondareva, V.M., Leonard, J.B., Plisetskaya, E.M., 2001. Neuropeptide
1114   regulation of feeding in catfish, Ictalurus punctatus: a role for glucagon-like peptide-1 (GLP-
1115   1)? Comp Biochem Physiol B Biochem Mol Biol 129, 623-631.
1116   Smith, J.J., Kuraku, S., Holt, C., Sauka-Spengler, T., Jiang, N., Campbell, M.S., Yandell, M.D.,
1117   Manousaki, T., Meyer, A., Bloom, O.E., Morgan, J.R., Buxbaum, J.D., Sachidanandam, R.,
1118   Sims, C., Garruss, A.S., Cook, M., Krumlauf, R., Wiedemann, L.M., Sower, S.A., Decatur,
1119   W.A., Hall, J.A., Amemiya, C.T., Saha, N.R., Buckley, K.M., Rast, J.P., Das, S., Hirano, M.,
1120   McCurley, N., Guo, P., Rohner, N., Tabin, C.J., Piccinelli, P., Elgar, G., Ruffier, M., Aken,
1121   B.L., Searle, S.M., Muffato, M., Pignatelli, M., Herrero, J., Jones, M., Brown, C.T., Chung-
1122   Davidson, Y.W., Nanlohy, K.G., Libants, S.V., Yeh, C.Y., McCauley, D.W., Langeland, J.A.,
1123   Pancer, Z., Fritzsch, B., de Jong, P.J., Zhu, B., Fulton, L.L., Theising, B., Flicek, P., Bronner,
1124   M.E., Warren, W.C., Clifton, S.W., Wilson, R.K., Li, W., 2013. Sequencing of the sea lamprey
1125   (Petromyzon marinus) genome provides insights into vertebrate evolution. Nat Genet 45, 415-
1126   421, 421e411-412.
1127   Southey, B.R., Amare, A., Zimmerman, T.A., Rodriguez-Zas, S.L., Sweedler, J.V., 2006.
1128   NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses
1129   of the resulting peptides. Nucleic Acids Res 34, W267-272.
1130   Takeda, J., Seino, Y., Tanaka, K., Fukumoto, H., Kayano, T., Takahashi, H., Mitani, T.,
1131   Kurono, M., Suzuki, T., Tobe, T., et al., 1987. Sequence of an intestinal cDNA encoding human
1132   gastric inhibitory polypeptide precursor. Proc Natl Acad Sci U S A 84, 7005-7008.
1133   Tine, M., Kuhl, H., Gagnaire, P.A., Louro, B., Desmarais, E., Martins, R.S., Hecht, J., Knaust,
1134   F., Belkhir, K., Klages, S., Dieterich, R., Stueber, K., Piferrer, F., Guinand, B., Bierne, N.,
1135   Volckaert, F.A., Bargelloni, L., Power, D.M., Bonhomme, F., Canario, A.V., Reinhardt, R.,
1136   2014. European sea bass genome and its variation provide insights into adaptation to
1137   euryhalinity and speciation. Nat Commun 5, 5770.
1138   Tseng, C.C., Jarboe, L.A., Landau, S.B., Williams, E.K., Wolfe, M.M., 1993. Glucose-
1139   dependent insulinotropic peptide: structure of the precursor and tissue-specific expression in
1140   rat. Proc Natl Acad Sci U S A 90, 1992-1996.
1141   Venkatesh, B., 2003. Evolution and diversity of fish genomes. Curr Opin Genet Dev 13, 588-
1142   592.
1143   Venkatesh, B., Lee, A.P., Ravi, V., Maurya, A.K., Lian, M.M., Swann, J.B., Ohta, Y., Flajnik,
1144   M.F., Sutoh, Y., Kasahara, M., Hoon, S., Gangu, V., Roy, S.W., Irimia, M., Korzh, V.,
1145   Kondrychyn, I., Lim, Z.W., Tay, B.H., Tohari, S., Kong, K.W., Ho, S.F., Lorente-Galdos, B.,
1146   Quilez, J., Marques-Bonet, T., Raney, B.J., Ingham, P.W., Tay, A., Hillier, L.W., Minx, P.,

1147 Boehm, T., Wilson, R.K., Brenner, S., Warren, W.C., 2014. Elephant shark genome provides
1148 unique insights into gnathostome evolution (vol 505, pg 174, 2014). Nature 513.
1149 Volkoff, H., Canosa, L.F., Unniappan, S., Cerda-Reverter, J.M., Bernier, N.J., Kelly, S.P.,
1150 Peter, R.E., 2005. Neuropeptides and the control of food intake in fish. Gen Comp Endocrinol
1151 142, 3-19.
1152 Volkoff, H., Peter, R.E., 2006. Feeding behavior of fish and its control. Zebrafish 3, 131-140.
1153 Wang, Y., Meng, F., Zhong, Y., Huang, G., Li, J., 2012. Discovery of a novel glucagon-like
1154 peptide (GCGL) and its receptor (GCGLR) in chickens: evidence for the existence of GCGL
1155 and GCGLR genes in nonmammalian vertebrates. Endocrinology 153, 5247-5260.
1156 Wang, Y., Nielsen, P.F., Youson, J.H., Potter, I.C., Conlon, J.M., 1999. Multiple forms of
1157 glucagon and somatostatin isolated from the intestine of the southern-hemisphere lamprey
1158 Geotria australis. Gen Comp Endocrinol 113, 274-282.
1159 Wei, Y., Mojsov, S., 1995. Tissue-specific expression of the human receptor for glucagon-like
1160 peptide-I: brain, heart and pancreatic forms have the same deduced amino acid sequences.
1161 FEBS Lett 358, 219-224.
1162 Wettergren, A., Schjoldager, B., Mortensen, P.E., Myhre, J., Christiansen, J., Holst, J.J., 1993.
1163 Truncated GLP-1 (proglucagon 78-107-amide) inhibits gastric and pancreatic functions in
1164 man. Dig Dis Sci 38, 665-673.
1165 Yeung, C.M., Wong, C.K., Chung, S.K., Chung, S.S., Chow, B.K., 1999. Glucose-dependent
1166 insulinotropic polypeptide gene expression in the stomach: revealed by a transgenic mouse
1167 study, in situ hybridization and immunohistochemical staining. Mol Cell Endocrinol 154, 161-
1168 170.
1169 Youson, J.H., Al-Mahrouki, A.A., 1999. Ontogenetic and phylogenetic development of the
1170 endocrine pancreas (islet organ) in fish. Gen Comp Endocrinol 116, 303-335.
1171 Yue, S., Irwin, D.M., 2005. Structure and expression of the chicken proglucagon gene. Mol
1172 Cell Endocrinol 230, 69-76.
1173 Zhou, L., Irwin, D.M., 2004. Fish proglucagon genes have differing coding potential. Comp
1174 Biochem Physiol B Biochem Mol Biol 137, 255-264.
1175
1176

Table 1

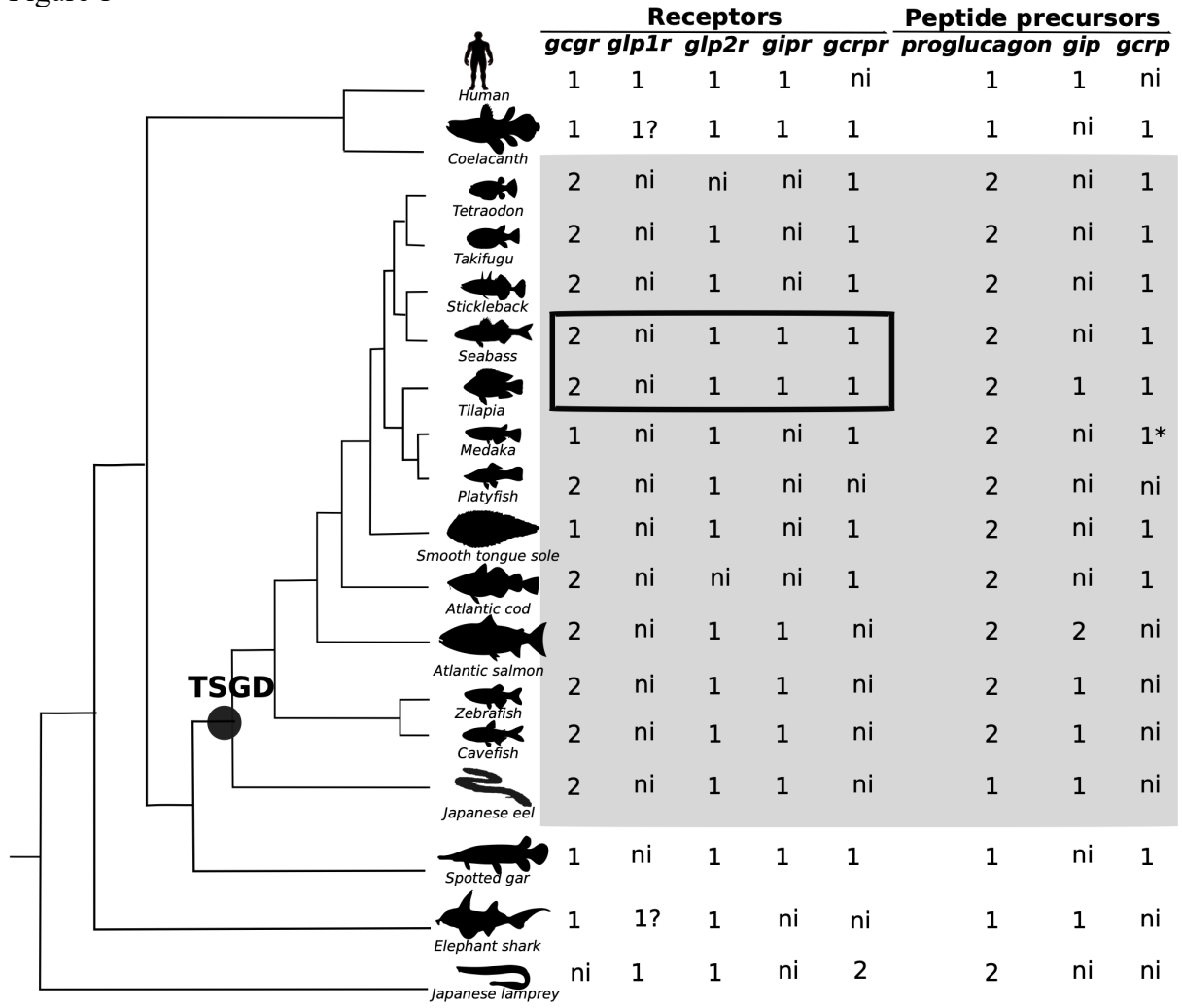|	Gcg	Glp1	*Proglucagon*	References
Ray-finned fish				
Alligator gar (*Lepisosteus spatula*)	Pancreas (29)	Pancreas (34)		(Pollock et al., 1988)
American eel (*Anguilla rostrata*)		Pancreas (30)		(Conlon et al., 1991)
European eel (*Anguilla anguilla*)	Pancreas (29, 35)			(Conlon et al., 1988)
Anglerfish (*Lophius americanus*)	Pancreas (29)	Pancreas (31)	Pancreas (*a* and *b*)	(Lund et al., 1982; Lund et al., 1983; Andrews et al., 1986; Nichols et al., 1988)
Coho salmon (*Oncorhynchus kisutch*)	Pancreas (29)	Pancreas (31)		(Plisetskaya et al., 1986)
Channel catfish (*Ictalurus punctata*)	Pancreas (29)	Pancreas (34)		(Andrews and Ronner, 1985)
Daddy sculpin (*Cottus scorpius*)	Pancreas (29)	Pancreas (31)		(Conlon et al., 1987b; Cutfield and Cutfield, 1993)
Pacific ratfish (*Hydrolagus colliei*)		Pancreas (35)		(Conlon et al., 1989)
Bowfin (*Amia calva*)	Pancreas (29)	Pancreas (34)		(Conlon et al., 1993b)
Bigeye tuna (*Thunnus obesus*)	Pancreas (29)			(Navarro et al., 1991)
Flounder (*Platichthys jlesus*)	Pancreas (29)			(Conlon et al., 1987a)
Paddlefish (*Polyodon spathula*)	Pancreas (29)	Pancreas (30)		(Nguyen et al., 1994)
Kaluga sturgeon (*Huso dauricus*)	Pancreas (29 34 and 35)			(Andoh et al., 2000)
Small-scaled pacu (*Piaractus mesopotamicus*)	Pancreas (29)	Pancreas (34)		(de Lima et al., 1999)
Rainbow trout (*Oncorhynchus mykiss*)			Intestine (*a*) Pancreas (*b*)	(Irwin and Wong, 1995)
Goldfish (*Carassius auratus*)			Intestine	(Yuen et al., 1997)
Ya-fish (*Schizothorax prenanti*)			Intestine	(Lin et al.)
Copper rockfish (*Sebastes caurinus*)	Pancreas (29)	Pancreas (31 and 34)	Intestine and brain (*a* and *b*)	(Busby and Mommsen, 2016)
Cartilaginous				
Small-spotted catshark (*Scyliorhinus canicula*)	Pancreas (29, 33) Intestine (29)	Pancreas (29)		(Conlon et al., 1987c; Conlon et al., 1994)
Ray (*Torpedo marmorata*)	Pancreas (29)			(Conlon and Thim, 1985)
Elephant shark (*Callorhynchus milii*)	Pancreas (29)			(Berks et al., 1989)
Agnatha				
Pouched lamprey (*Geotria australis*)	Intestine (29) I and II			(Andoh et al., 2000)
Sea lamprey (*Petromyzon marinus*)	Intestine (29)	Intestine (32)	Intestine I and II	(Conlon et al., 1993a; Irwin et al., 1999)
River lamprey (*Lampetra fluviatilis*)	Intestine (29)			(Conlon et al., 1995)

1180   Table 2
1181

Table 1: Primer sequences used to amplify the glucagon-like receptors and precursors

Primer	Sequence (5'-3')	Tm (ºC)	Eff. (%)	R²
Receptors				
*gcgra*fwd	acggttagctgtcgcatcg	60	99.1	0.99
*gcgra*rev	accaggaggttgtggaggtaga			
*gcgrb*fwd	tgtttccctgcaaaaggtgac	62	98.6	0.99
*gcgrb*rev	gcacaacaacagtagtgttagg			
*glp2r*fwd	caccaggtatcaggctcaga	62	96.5	1.00
*glp2r*rev	cattcccaggagatgaatgg			
*gipr*fwd	ctcgtttcagggtttacttgta	60	96.3	0.99
*gipr*rev	gagcggcagacagggaca			
*gcrpr*fwd	ctgctgtttctcaaactgccaa	62	105.8	0.99
*gcrpr*rev	tgttgaccacagagctggca			
Precursors				
*progcga*fwd	ctacctcaaggaccaggcaat	62	99.1	0.99
*progcga*rev	ctgacgacttggaggtcatga			
*progcgb*fwd	tatgctgaccgaacctattgag	60	104.5	0.99
*progcgb*rev	cgtctgcatggcgtctaaata			
*gcrp*fwd	cattcagatgggacattcaca	-	-	-
*gcrp*rev	tggtgctggccagccac			
*gip*1fwd	tgttgttcaaacttctgatcatc	-	-	-
*gip*1rev	gcaggaacttataaagtttttctg			
*gip*2fwd	tccaccatcgccagcga	-	-	-
*gip*2rev	agcgggcttggttttcttct			
Reference genes				
*Ef1alpha*fwd	gacacagagacttcatcaag	58	98.5	1.00
*Ef1alpha*rev	gtccgttcttagagatacca			
*18S*fwd	tgacggaagggcaccaccag	58	98.3	0.99
*18S*rev	aatcgctccaccaactaagaacgc			

1182

Figure 1

	Receptors					Peptide precursors		
	gcgr	glp1r	glp2r	gipr	gcrpr	proglucagon	gip	gcrp
Human	1	1	1	1	ni	1	1	ni
Coelacanth	1	1?	1	1	1	1	ni	1
Tetraodon	2	ni	ni	ni	1	2	ni	1
Takifugu	2	ni	1	ni	1	2	ni	1
Stickleback	2	ni	1	ni	1	2	ni	1
Seabass	2	ni	1	1	1	2	ni	1
Tilapia	2	ni	1	1	1	2	1	1
Medaka	1	ni	1	ni	1	2	ni	1*
Platyfish	2	ni	1	ni	ni	2	ni	ni
Smooth tongue sole	1	ni	1	ni	1	2	ni	1
Atlantic cod	2	ni	ni	ni	1	2	ni	1
Atlantic salmon	2	ni	1	1	ni	2	2	ni
Zebrafish	2	ni	1	1	ni	2	1	ni
Cavefish	2	ni	1	1	ni	2	1	ni
Japanese eel	2	ni	1	1	ni	1	1	ni
Spotted gar	1	ni	1	1	1	1	ni	1
Elephant shark	1	1?	1	ni	ni	1	1	ni
Japanese lamprey	ni	1	1	ni	2	2	ni	ni

TSGD

1186   Figure 2
1187

```
Human          ------------------------------------------------------
Coelacanth     ------------------------------------------------------
Spotted gar    ------------------------------------------------------
Elephant shark MQISFRFQQSFGPREATVCPAISFIQGSSAAGNTMKGVDSISVLLLLILV


Human          ----MKSIYFVAG-LFVMLVQGSWQRSLQDTEEKSRSFSASQADPLSDPD
Coelacanth     ----MKTTHSVAGIIVLMLIQGSWQNPLQDIENKSRLFKAANTEPIDEPR
Spotted        ----MKGIHSLAGLLLLIIVQGSWQVPLQDTEDNSRLLTED--SSIDEPR
Elephant       QNTTMKGVDSISVLLLLILVQNTWQTPIQDS-DSSRALETAEKQPVVTPN
                   **     ::   :.::::*.:** .:**  :.** :     ..:   *
```

Gcg
```
Human          QMNEDKRHSQGTFTSDYSKYLDSRRAQDFVQWLMNTKRNRNN--------
Coelacanth     ELTEVKRHSQGTFTSDYTKYLDTIRAQDFVQWLMSTKRSG----------
Spotted gar    ELTNVKRHSQGTFTNDYSKYLDTRRAQDFVQWLMSTKRSGG---------
Elephant shark VMTDVKRHSEGTFSSDYSKYLDSRRAKDFVQWLMSTKRNGANTDKTKRHA
                .: ****:***:.**:****: **:*******.***.
```

Glp-like
```
Human          -------------------------------IAKRHDEFERHAEGT
Coelacanth     ----------------------------------------------
Spotted gar    ----------------------------ITRRHADGT
Elephant shark DGSYTSDVESLSDYFKAKRFVDSLTNYNKHQSDRRMSKRNADRASHTEED
                                                      *:::
```

Glp1
```
Human          FTSDVSSYLEGQAAKEFIAWLVKGRGRRDFPEEVA------IVEELGRRH
Coelacanth     ------------------------FPNEIS------ETEGMDRRH
Spotted gar    YTSDVSSYLQDQAAKKFVTWLKQGQDRRDFSEESS------ETEEMYRRH
Elephant shark YPSDFSSYLEAKAARDFINWLIKGRGRRDFAEESREIENEVIAEELDRRH
                :: ** ****: :**:.*:*** :**.** *.:*        .* : ***
```

Glp2
```
Human          ADGSFSDEMNTILDNLAARDFINWLIQTKITD-RK-----------
Coelacanth     ADGSFTSDINKVLDTIAAKEFLNWLINSKDSQPRDPFSENQ------
Spotted gar    ADGSFTSDVNNVLDSIAAKEFIIWVMNSKPSEESEPFNV--------
Elephant shark ADGTFTSEINVVLDTMAAKEFLNWIINSKNIQSRDAEKDFFNEYKGR
                ***:*:.::* :**.:**::*: *::::*  :
```
1188
1189
1190
```

Figure 3

Figure 4

1199    Figure 5

Figure 6

1207    Figure 7



**17**    50.42  12.78  36.60    40.13  48.95  48.92  48.83  48.76    48.13  48.07  48.04

**JH126617**    1.59  1.56    0.70  0.53  0.52

*Human*

**12**    48.80  57.60  56.66  56.44  57.09  51.84    51.59  50.84  50.50  53.50  48.14  50.39

**JH126563**    3.12  4.05  6.46    5.41  5.29  5.14    7.29    7.46  7.38

*Coelacanth*

**LG1B**    15.69  15.61  15.44  15.42  15.42  15.41    15.41  15.39  15.38    15.31  15.30  15.29

*adam8b  acsf2  arhgap44  mrm1  bc2  msl1  glp  snf8  calcoco2  ttll6  hoxb  skap1  cbx1b  nfe2l1b*

**GL831200**    1.71  1.81  1.88  2.07  2.07  2.08  2.08  2.09  2.09  2.11  2.16  2.21  2.22

*Seabass*

*hoxc  cacnb3b  dtx3  ptges3a  mlpa  stat6  fignl2  gcrp  scn8  tmprss12  dlp2ba  atf7  asb8  larp4aa*

**LG22-25**    13.17  13.21  13.22  13.23  13.24  13.26  13.31  13.32  13.36  13.37  13.41  13.43  13.44

**19**    17.94  17.72  17.63  17.63  17.62  17.61  17.60  17.52  17.51

*Tilapia*

**GL831196**    0.62  0.64  0.56  0.54  0.52  0.48  0.43  0.42  0.38  0.35  0.35  0.33

**12**    9.05  26.40  48.49  26.93  26.93  20.54  20.53  20.52  27.04  27.06  28.75  28.73

*Medaka*

**7**    11.73  19.38  11.60  11.54  11.51  11.49  11.46  11.42  11.41  11.39

**LG15**    9.50  9.38  9.31  8.94  8.89  8.87

*Zebrafish*

**23**    27.02  27.11  27.12  27.14  27.15  23.22  27.42  27.43  27.52  27.57  27.58

*Spotted gar*

**LG4**    18.16  18.07  18.02  16.43  16.44  16.57  16.64  16.67  16.76  16.84  16.87  17.70

**sc_122**    1.00  1.02  1.46  1.74

*Elephant shark*

1208
1209
1210

1214　Figure 9

1215



1216
1217
1218

1219    Figure 10

Figure 11

## Supplementary Table 1

| | gcgr | glp1r | glp2r | gipr | gcrp | gcgr-like |
|---|---|---|---|---|---|---|
| **Tetrapod** | | | | | | |
| Human (*Homo sapiens*) | NP_000151.1 | AAR05444.1 | AAH96263.1 | AAC97984.1 | ni | |
| Mouse (*Mus musculus*) | NP_032127.2 | NP_067307.2 | NP_783612.2 | NP_001074284.1 | ni | |
| Chicken (*Gallus gallus*) | ENSGALG00000011219 | ENSGALG00000010078 | ENSGALG00000041800 | ni | ACH90400.1 | |
| Lizard (*Anolis carolinensis*) | ENSACAG00000010639 | ENSACAG00000001843 | ENSACAG00000016945 | ENSACAG00000009350 | ENSACAG00000001229 | |
| Chinese softshell turtle (*Pelodiscus sinensis*) | n.i. | ENSPSIG00000004863 | ENSPSIG00000018008 | ENSPSIG00000009150 | ENSPSIG00000017703 | |
| Clawed frog (*Xenopus tropicalis*) | ENSXETG00000005421 | ENSXETG00000032385 | ENSXETG00000016321 | ENSXETG00000018255 | ENSXETG00000010322 | |
| **Lobe-finned fish** | | | | | | |
| Coelacanth (*Latimeria chalumnae*) | ENSLACG00000014981 | ENSLACG00000014381? | ENSLACG00000003010 | ENSLACG00000007720 | ENSLACG00000003192 | |
| | | | | | n.i. | |
| **Ray-finned fish** | | | | | | |
| Takifugu (*Takifugu rubripes*) | ENSTRUG00000012066 ENSTRUG00000014346 | n.i. | ENSTRUG00000017202 | n.i. | ENSTRUG00000015257 | |
| Tetraodon (*Tetraodon nigroviridis*) | ENSTNIG00000010439 ENSTNIG00000012358 | n.i. | n.i. | n.i. | ENSTNIG00000012820 | |
| Nile tilapia (*Oreochromis niloticus*) | ENSONIG00000001985 ENSONIG00000016459 | n.i. | ENSONIG00000019565 | XP_003459508.2 (GL831556.1: 0,23 Mb) | GL831136.1: 8,195 Mb | |
| Sea bass (*Dicentrarchus labrax*) | DLAgn_00191040 DLAgn_00178590 | n.i. | DLAgn_00240050 | DLAgn_00036000 | DLAgn_00100590 | |
| Medaka (*Oryzias latipes*) | ENSORLG00000007082 | n.i. | ENSORLG00000013155 | n.i. | ENSORLG00000007568 | |
| Amazon molly (*Poecilia formosa*) | ENSPFOG00000002015 ENSPFOG00000018824 | n.i. | ENSPFOG00000018904 | n.i. | ENSPFOG00000015868 | |
| Stickleback (*Gasterosteus aculeatus*) | ENSGACG00000010535 ENSGACG00000018868 | n.i. | ENSGACG00000014797 | n.i. | ENSGACG00000009099 | |
| Platyfish (*Xiphophorus maculatus*) | ENSXMAG00000012984 ENSXMAG00000012523 | n.i. | ENSXMAG00000005305 | n.i. | n.i. | |
| Cod (*Gadus morhua*) | ENSGMOG00000008208 ENSGMOG00000008506 | n.i. | ENSGMOG00000005923 | n.i. | ENSGMOG00000006250 | |
| Smooth tongue sole (*Cynoglossus semilaevis*) | XP_016895711.1 | n.i. | XP_008313439.1 | n.i. | XP_008312440.1 | |
| Salmon (*Salmon salar*) | XP_014038281.1 XP_014049938.1 | n.i. | XP_014035029.1 | XP_014069579.1 | n.i. | |
| Zebrafish (*Danio rerio*) | ENSDARG00000104022 ENSDARG00000036272 | n.i. | XP_009304634.1 (Chr 12: 0,316 Mb) | ENSDARG00000025478 | n.i. | |
| Cavefish (*Astyanax mexicanus*) | ENSAMXG00000017308 ENSAMXG00000007858 | n.i. | ENSAMXG00000016128 | ENSAMXG00000010353 | n.i. | |
| Japanese eel (*Anguilla japonica*) | KI305549.1 (scaffold1162) AVPY01196538.1 (scaffold5045) | n.i. | KI304458.1 (scaffold71) | AVPY01197693.1 (scaffold5094) | n.i. | |
| Spotted gar (*Lepisosteus oculatus*) | ENSLOCG00000013910 | n.i. | ENSLOCG00000011822 | ENSLOCG00000014786 | ENSLOCG00000003380 | |
| **Cartilaginous fish** | | | | | | |
| Elephant shark (*Callorhinchus milii*) | SINCAMG00000002895 | SINCAMG00000015132? | SINCAMG00000004406 | n.i. | n.i. | |
| **Jawless fish** | | | | | | |
| Sea lamprey (*Petromyzon marinus*) | n.i. | n.i. | ENSPMAG00000003877 | n.i. | ENSPMAG00000008188 ENSPMAG00000003562 | |
| Japanese lamprey (Lethenteron japonicum) | n.i. | JL7967 | JL5602 | n.i. | JL10168 JL9019 | |
| **Tunicate** | | | | | | |
| Ciona | | | | | | ENSCING00000006559 ENSCING00000006557 |

1227

Supplementary Table 2

| | *Proglucagon* | *gip* | *gcrp* |
|---|---|---|---|
| **Tetrapod** | | | |
| Human (*Homo sapiens*) | AAH05278.1 | AAH96149.1 | n.i. |
| Mouse (*Mus musculus*) | AAH12975.1 | EDL16005.1 | n.i. |
| Chicken (*Gallus gallus*) | NP_001177094 | ABL10366.1 | ENSGALG00000042837 |
| Lizard (*Anolis carolinensis*) | ENSACAG00000014182+ | n.i. | ENSACAG00000028018 |
| Chinese softshell turtle (*Pelodiscus sinensis*) | ENSPSIG00000010712 | ENSPSIG00000005451 | n.i. |
| Clawed frog (*Xenopus tropicalis*) | ENSXETG00000013178 | ENSXETG00000031431 | ENSXETG00000008236 |
| **Lobe-finned fish** | | | |
| Coelacanth (*Latimeria chalumnae*) | XP_006004407.1 | n.i. | JH126563.1: 5,29 Mb |
| **Ray-finned fish** | | | |
| Takifugu (*Takifugu rubripes*) | ENSTRUG00000008721 ENSTRUG00000004633 | n.i. | scaffold_66: 1,03 Mb |
| Tetraodon (*Tetraodon nigroviridis*) | ENSTNIG00000013278 ENSTNIG00000000614 | n.i. | Chr9: 3,34 Mb |
| Nile tilapia (*Oreochromis niloticus*) | ENSONIG00000018307 ENSONIG00000008919 | GR702825* | GL831196.1: 0,48Mb |
| Sea bass (*Dicentrarchus labrax*) | DLAgn_00138270 DLAgn_00050390 | n.i. | LG22-25:13,31 Mb |
| Medaka (*Oryzias latipes*) | ENSORLG00000002782 ENSORLG00000016891 | n.i. | Chr21: 23,11 Mb |
| Amazon molly (*Poecilia formosa*) | ENSPFOG00000007857 ENSPFOG00000017328 | n.i. | KI519690: 2,02 Mb |
| Stickleback (*Gasterosteus aculeatus*) | ENSGACG00000013877 ENSGACG00000005606 | n.i. | groupI: 20,20 Mb |
| Platyfish (*Xiphophorus maculatus*) | ENSXMAG00000013819 ENSXMAG00000011481 | n.i. | n.i. |
| Cod (*Gadus morhua*) | ENSGMOG00000003985 ENSGMOG00000014909 | n.i. | GeneScaffold_2603: 0,39 Mb |
| Smooth tongue sole (*Cynoglossus semilaevis*) | XP_008323075.2 XP_008326926.1 | n.i. | NC_024316.1 |
| Salmon (*Salmon salar*) | SS2U042201 SS2U025187 | ABW77503.1 XP_014014591.1 | n.i. |
| Zebrafish (*Danio rerio*) | ENSDARG00000079296 ENSDARG00000040907 | AAI46706.1 | n.i. |
| Cavefish (*Astyanax mexicanus*) | ENSAMXG00000013524 ENSAMXG00000010007 | ENSXMAG00000011481 | n.i. |
| Japanese eel (*Anguilla japonica*) | KI305999.1(scaffold_1612) KI314819.1(scaffold_10445) | KI307715.1(scaffold_3328) | n.i. |
| Spotted gar (*Lepisosteus oculatus*) | ENSLOCG00000008502 | n.i. | LG4: 16,6 Mb |
| **Cartilaginous fish** | | | |
| Elephant shark (*Callorhinchus milii*) | SINCAMG00000000174 | SINCAMG00000003496 | n.i. |
| **Jawless fish** | | | |
| Sea lamprey (*Petromyzon marinus*) | ENSPMAG00000002186 ENSPMAG00000005961 | n.i. | n.i. |

1228 + only gcg found

1233    Supplementary Figure 2



1234
1235

1236 Supplementary Figure 3

1237



1238

1239 Supplementary Figure 4

1240



1241
1242
1243
1244

1245 Supplementary File
1246
1247 >Human_NP_000151.1_GCGR
1248 VMYTVGYSLSLGALLLALAILGGLSKLHCTRNAIHANLFASFVLKASSVLVIDGLLRTRYSQKIGDDLSVSTWLS
1249 DGAVAGCRVAAVFMQYGIVANYCWLLVEGLYLHNLLGLATLPERSFFSLYLGIGWGAPMLFVVPWAVVKCLFENV
1250 QCWTSNDNMGFWWILRFPVFLAILINFFIFVRIVQLLVAKLRARQMHHTDYKFRLAKSTLTLIPLLGVHEVVFAF
1251 VTDEHAQGTLRSAKLFFDLFLSSFQGLLVAVLYCFLN
1252
1253 >Human_AAR05444.1_GLP1R
1254 IIYTVGYALSFSALVIASAILLGFRHLHCTRNYIHLNLFASFILRALSVFIKDAALKWMYSTAAQQHQWDGLLSY
1255 QDSLSCRLVFLLMQYCVAANYYWLLVEGVYLYTLLAFSVFSEQWIFRLYVSIGWGVPLLFVVPWGIVKYLYEDEG
1256 CWTRNSNMNYWLIIRLPILFAIGVNFLIFVRVICIVVSKLKANLMCKTDIKCRLAKSTLTLIPLLGTHEVIFAFV
1257 MDEHARGTLRFIKLFTELSFTSFQGLMVAILYCFVN
1258
1259 >Human_AAH96263.1_GLP2R
1260 LMYTVGYSFSLISLFLTLTLLLFLRKLHCTRNYIHMNLFASFILRTLAVLVKDVVFYNSYSKRPDNENGWMSYLS
1261 EMSTSCRSVQVLLHYFVGANYLWLLVEGLYLHTLLEPTVLPERRLWPRYLLLGWAFPVLFVVPWGFARAHLENTG
1262 CWTTNGNKKIWWIIRGPMMLCVTVNFFIFLKILKLLISKLKAHQMCFRDYKYRLAKSTLVLIPLLGVHEILFSFI
1263 TDDQVEGFAKLIRLFIQLTLSSFHGFLVALQYGFAN
1264
1265 >Human_AAC97984.1_GIPR
1266 VMYTVGYSLSLATLLLALLILSLFRRLHCTRNYIHINLFTSFMLRAAAILSRDRLLPRPGPYLGDQALALWNQAL
1267 AACRTAQIVTQYCVGANYTWLLVEGVYLHSLLVLVGGSEEGHFRYYLLLGWGAPALFVIPWVIVRYLYENTQCWE
1268 RNEVKAIWWIIRTPILMTILINFLIFIRILGILLSKLRTRQMRCRDYRLRLARSTLTLVPLLGVHEVVFAPVTEE
1269 QARGALRFAKLGFEIFLSSFQGFLVSVLYCFIN
1270
1271 >Human_AAI12248_1_PTH1R
1272 MIYTVGYSVSLASLTVAVLILAYFRRLHCTRNYIHMHLFLSFMLRAVSIFVKDAVLYSGATLDEAERLTEEELRA
1273 IAQAPPPPATAAAGYAGCRVAVTFFLYFLATNYYWILVEGLYLHSLIFMAFFSEKKYLWGFTVFGWGLPAVFVAV
1274 WVSVRATLANTGCWDLSSGNKKWIIQVPILASIVLNFILFINIVRVLATKLRETNAGRCDTRQQYRKLLKSTLVL
1275 MPLFGVHYIVFMATPYTEVSGTLWQVQMHYEMLFNSFQGFFVAIIYCFCN
1276
1277 >Mouse_NP_032127.2_gcgr
1278 VMYTVGYSLSLGALLLALVILLGLRKLHCTRNYIHGNLFASFVLKAGSVLVIDWLLKTRYSQKIGDDLSVSVWLS
1279 DGAMAGCRVATVIMQYGIIANYCWLLVEGVYLYSLLSLATFSERSFFSLYLGIGWGAPLLFVIPWVVVKCLFENV
1280 QCWTSNDNMGFWWILRIPVFLALLINFFIFVHIIHLLVAKLRAHQMHYADYKFRLARSTLTLIPLLGVHEVVFAF
1281 VTDEHAQGTLRSTKLFFDLFLSSFQGLLVAVLYCFLN
1282
1283 >Mouse_NP_067307.2_glp1r
1284 IIYTVGYALSFSALVIASAILVGFRHLHCTRNYIHLNLFASFILRALSVFIKDAALKWMYSTAAQQHQWDGLLSY
1285 QDSLGCRLVFLLMQYCVAANYYWLLVEGVYLYTLLAFSVFSEQRIFKLYLSIGWGVPLLFVIPWGIVKYLYEDEG
1286 CWTRNSNMNYWLIIRLPILFAIGVNFLIFIRVICIVVSKLKANLMCKTDIKCRLAKSTLTLIPLLGTHEVIFAFV
1287 MDEHARGTLRFIKLFTELSFTSFQGLMVAILYCFVN
1288
1289 > Mouse_NP_783612.2_glp2
1290 LMYTVGYSLSLISLFLALTLFLFLRKLHCTRNYIHMNLFASFILRALVVLVKDMVFYNSYSRRPDSESGWMSYLS
1291 EISASCRSVQVLLHYFVGTNHLWLLVEGLYHALLEPTVLPERRLWPKYLVVGWAFPMLFVIPWIFVRASLENTG
1292 CWAVNENKKIWWIIRGPILLCVTVNFFIFLKILKLLISKFRAHQMCFRDYKYRLAKSTLLLILLMGVHEFLFTFF
1293 TDDQVQGFSRLIRLFIQLTLSSFHGFLVALQYGFAS
1294
1295 >Mouse_NP_001074284.1_gipr
1296 IMYTVGYSLSLTTLLLALLILSLFRRLHCTRNYIHMNLFTSFMLRAAAILTRDQLLPPLGPYTGDQAPTPWNQAL
1297 AACRTAQIMTQYCVGANYTWLLVEGVYLHHLLVIVGRSEKGHFRCYLLLGWGAPALFVIPWVIVRYLRENTQCWE
1298 RNEVKAIWWIIRTPILITILINFLIFIRILGILVSKLRTRQMRCPDYRLRLARSTLTLVPLLGVHEVVFAPVTEE
1299 QVEGSLRFAKLAFEIFLSSFQGFLVSVLYCFIN
1300
1301 >Chicken_ENSGALG00000011219_gcgr
1302 VMYTVGYSVSLCALLLALAVLLGFSKLHCMRNYIHMNLFASFIVKGVSVLVIDALLKTHYSDKIDDYNVRIWLSD
1303 EAAAGCRAATVFMQYGIVANYCWLLVEGIYLHNLLVVAVFSERSYFTLYLCIGWGAPVLFLIPWVVVKFLYENIQ
1304 CWSTNHNMGFWWILRFPVFLAILINFFIFIRIIQILVSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHEVIFAFI
1305 TDEHAQGTLRYVKLFFDLFLSSFQGMLVAILYCFVN

```
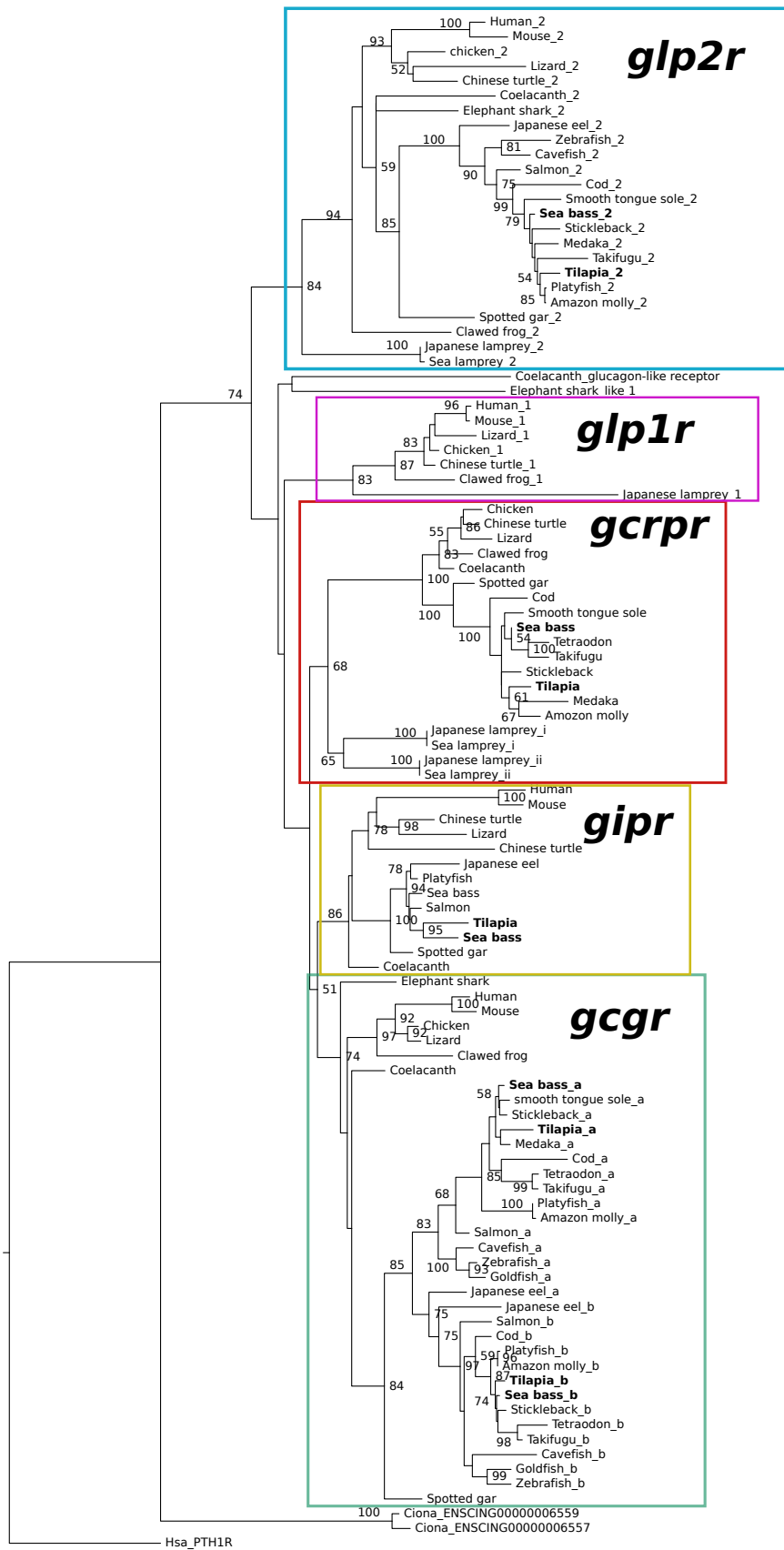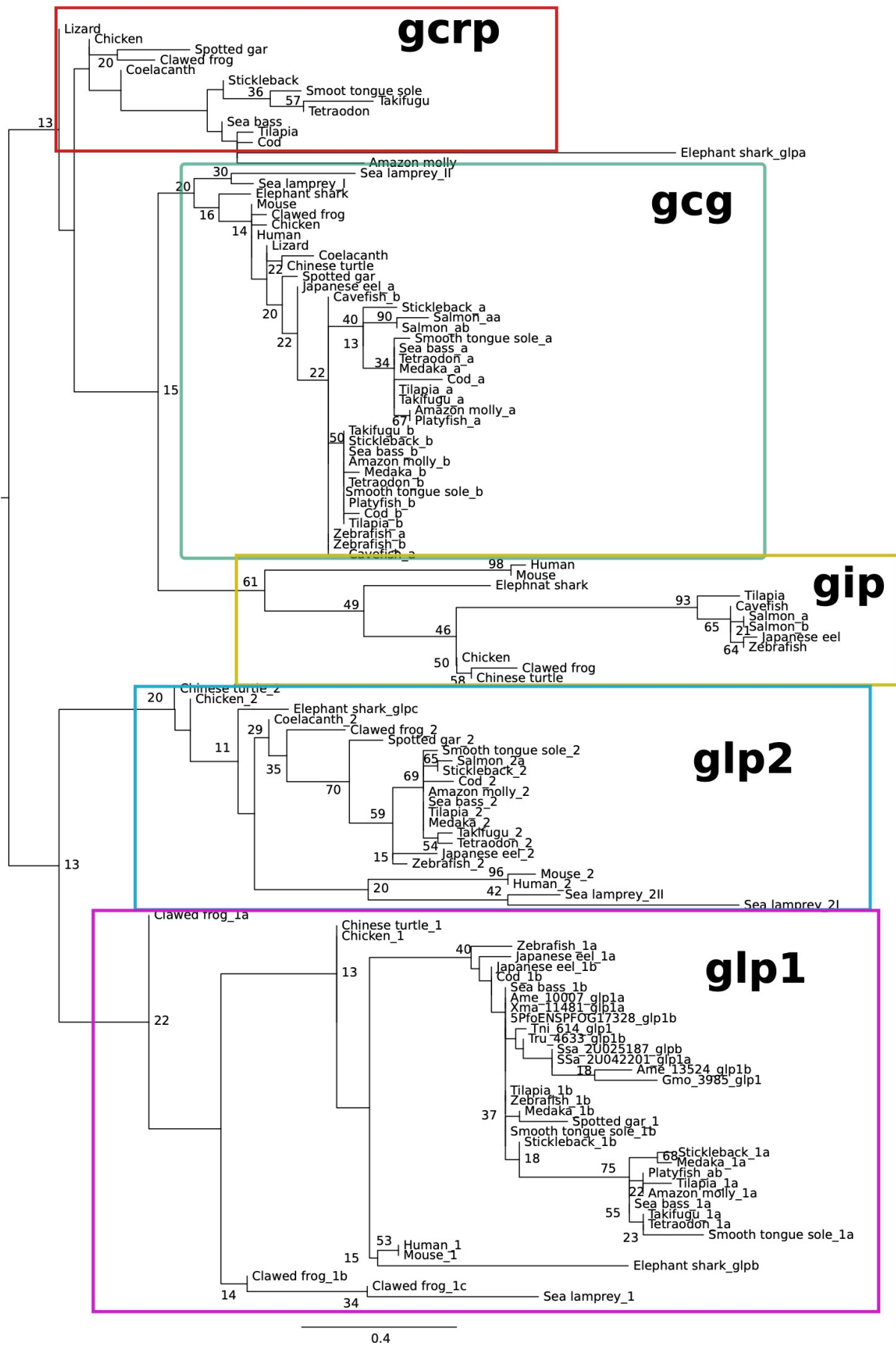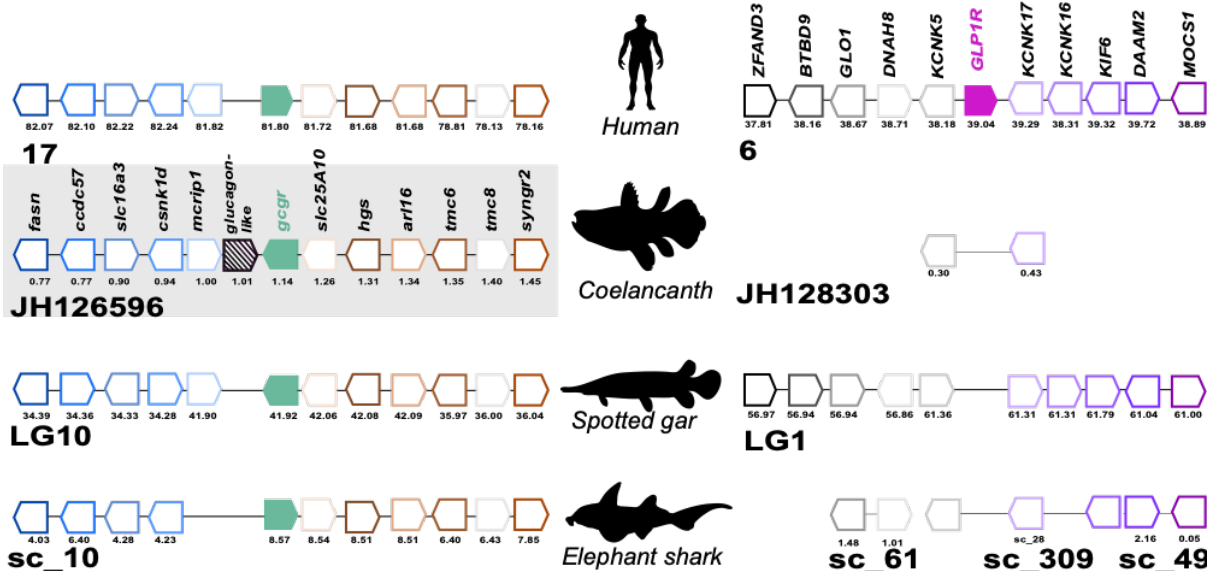1306
1307    >Chicken_ENSGALG00000010078_glp1r
1308    IIYTIGYALSFSALVIATAILLGFRHLHCTRNYIHLNLFTSFILRAISVFIKDSVVKWMYSTATQEHQWEGLISF
1309    QESLSCRLVFVMMQYCVAANYYWLLVEGMYLYTLLVLSVFSEQRIFRLYLCIGWGVPMLFVILWGTVKYLYEDEG
1310    CWSRNYNMNYWLIIRLPILIAIGVNFLIFIRVICIIISKLQANLMCKTDIKCRLAKSTLTLIPLLGTHEVIFAFI
1311    TDEHARGMLRFVKLFTELSFASFQGLMVAILYCFIN
1312
1313    >Chicken_ENSGALG00000041800_glp2r
1314    LLYTIGYCFSLISLVLALLILSLLRKLHCTRNYIHMNLFASFILRATAILIKDTVLHKIYSKRPNDETGWILYLS
1315    PEILIICRAAQFFMHYFVGANYFWLLVEGIYLHTLLITAVLSERRLLQTYIVIGWAVPILFVAPWGISRSKLENT
1316    GCWGTNEHMGIWWIIRGPMLFSITVNFGIFLKILRMLISKLKAQQMSFHDYKYRLARSTLVLIPLLGIHEFIFSF
1317    ITDEQVEGFLRHVRLFIQLTMSSFHGFFVAVLYCFAN
1318
1319    >Chicken_ACH90400.1_gcrpr
1320    VLYTVGYALSLLTLVSALLVLTVFRKLHCTRNYIHANLFASFGLRATSVMVKDALLERRWGAEVLQVADWQALLS
1321    HEAALGCRAAQVLMQYCILANHYWFLVEAVYLYKLLIGAVFSEKNYYRLYLYLGWGTPVVFVVPWMAAKYLKENA
1322    ECWALNENMAYWWIIRIPILLASMINLLIFMRILKVILAKLRANQKGYADYKLRLAKATLTLIPLFGIHEVVFIF
1323    ATDEQTTGILRYIKVFFTLFLNSFQGFLVAVLYCFAN
1324
1325    >Lizard_ENSACAG00000010639_gcr
1326    VMYTVGYSVSLCALLLALAVLLGFSKLHCMRNYIHMNLFASFILKGISVLIIDMLLNTRYSEKIDDYNVGLWLSH
1327    EAAAGCRAATVFMQYGIVANYCWLLVEGIYLHNLLVLAVFSERSYFTLYLFIGWGAPILFIVPWVVVKFLYENIQ
1328    CWSTNNNMGFWWILRFPVFLAIFINFFIFIRIIQILVSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHEVVFAFV
1329    TDEHAQGTLRYVKLFFDLFLSSFQGMLVAILYCFVN
1330
1331    >Lizard_ENSACAG00000001843_glp1r
1332    IIYTTGYALSFSALVIASGILLGFRHLHCTRNYIHLNLFASFILRAASIFIKDSMITWMYKTAPREEQWENLISY
1333    QESLSCRLIIVMMQYCVTANYYWLLVEGMYLYTLLALSVFSEQRIFRLYLCIGWGVPMLFVIFWGIVKYLYEDEG
1334    CWNKNLNMNYWLIIRLPILVAIGVNFLIFIRVICIIISKLQANLLRKTDIKCRLAKSTLTLIPLLGTHEIVFAFV
1335    TDEHAKGTLRFVKLFFELSFSSFQGLLVAILYCFNN
1336
1337    >Lizard_ENSACAG00000016945_glp2r
1338    NLYTVGYSFSFASLVLALIIMLPLRKLHCTRNYIHMNLFASFILRTIGVLIKDSVTHHTYFFIILDTEKPNDLNG
1339    WTSTLGSEMLALCRMAPLFMHYVVGANYFWLLVEGIYLHRLLTTVVSEKHQLVKYIFIGWGIPVLFVTSWGIIKY
1340    QLEHEGCWATHQNMAFWWIIRGPILFSILVNFIIFLNLLKLLHSKLKAQQMNCRDYKFRLARSTFVLISLLGIHE
1341    IVFSFITDEQIEGFYRHIRNFVQLTIGSFHGFLVALLYCFCN
1342
1343    >Lizard_ENSACAG00000001229_gcrpr
1344    VIYTVGYTLSLLALLLALLILTVFRKLRCTRNYIHANLFASFGLRAISVIVKDALLKKLWEREVFQVSEWETFLT
1345    HEAAIGCRVAQVVMQYCILANHYWFLVEAVYLYKLLIGAVFSEKNYYTLYLYLGWGTPVAFVVPWMAAKYLKENT
1346    ECWGENENMAYWWIIRIPILLASVINLLIFMQILKVILAKLRASQKSYADYKLRLAKATLTLVPLFGIHEVVFIF
1347    ATDEQTTGVLRYIKVFFTLFLNSFQGFLVAVLYCFAN
1348
1349    >Lizard_ENSACAG00000009350_gipr
1350    LMYTMGYSVSLMASLVALALLSTLRKLRCIRNYIHMNLFLSYMLRAISILTRDALLWLRFPEDFQKEGDFSSFPM
1351    GQAGASCRLAQVLTQYCVCANYYWLLVEGLYLHNLLGPLAFSEESYFPGYLLLGWGSPILFVIPWVIVRYLYENH
1352    ECWERNDNMGFWWIIRCPILLAIFVNLVIFIRIIKILVSKLQAQQMRYADYKMRLAKSTLTLIPLLGIHEVVFAL
1353    VTEEQAQGTLRYVKFFFELFLNSLQGLLVSILYCFIN
1354
1355    >Chinese softshell turtle_ENSPSIG00000004863_glp1r
1356    VIYTVGYALSFSALVIATAILLGFRHLHCTRNYIHLNLFTSFILRAVSVFIKDSVLKWMYSTATHEHQWEGLISY
1357    QESLSCRLVFVMMQYCVAANYYWLLVEGMYLHTLLVLSAISEQRVFRLYLCIGWGIPMLFVVLWGIVKYLYEDEG
1358    CWTRNHNMNYWLIIRLPILIAIGVNFLIFIRVICIIISKLQSNLMCKTDTKCRLAKSTLTLIPLLGTHEVIFAFV
1359    TDEHARGMLRIVKLFCELSFTSFQGLMVAILYCFIN
1360
1361    >Chinese softshell turtle_ENSPSIG00000018008_glp2r
1362    LLYTVGYYLSLSSLVLALLILLLLRKLHCTRNYIHINLFASFILRAIAVLIKDTITHNNYSKKPDNETGWISYFN
1363    PEILAICRTSQLSLHYFVGANYFWLLVEGIYLHTLLVTDVLTARLLLQRYIVIGWVVPVLFVAPWGIARAKLEDT
1364    GCWATHDNEGIWWIIRGPMLFSIAVNFCIFLKILKLLISKLKAQQMNFHDYKYRLARSTLVLIPLLGIQEFVFTF
1365    LPDEQVEGFSRRIRVFIQLIMSSFQGFLVAIFYCFVN
1366
```

```
1367    >Chinese softshell turtle_ENSPSIG00000009150_gipr
1368    LMDTVGYSLSLVSLLLALLLLLAFRKLRCTRNYIHANLFLSFMLRAVSILTRDALLRVRFSKGLLYEGDLFQLLG
1369    DQVPPGRERNSMWGAVPSAWGVKARGLLRSAAWQELERDWFCLRLSLPAGAPILFVVPWVIIRYLYENNQCWERN
1370    DNMAYWWIIRCPILLAILVNFVIFVRIIKILVSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHEVVFALVPEEQA
1371    QGTLRYIKLFFELFLSSLQGLLVSILYCFLN
1372
1373    >Chinese softshell turtle_ ENSPSIG00000017703_gcrpr
1374    VLYTVGYSLSLLALVSALLILTVFRKLRCTRNYIHANLFASFGLRAISVIARDTLLEKRWGMEILQVSDWEALLS
1375    NEGALGGRQVSEFQRPGRVGGGVRSSQRAEIRLRGRLGSGLTPKGFPVTWLSRGPRTPVVFVVPWLVAKYLKENA
1376    ECWALNENMAYWWIIRIPILLASLINLLIFMRILKVILAKLRANQKGYADYKLRLAKATLTLIPLFGIHEVVFIF
1377    ATDEQTTGILRYIKVFSMLFLNSFQGFLVAVLYCFAN
1378
1379    >Clawed frog_ENSXETG00000005421_gcgr
1380    RLSKLHCMRNYIHINLFASFILRAVSVLVLDTMLKTRYNENEKFEDTHLWLSSEALVGCRVAAVLMQYGIIANYY
1381    WLMVEGIYLYNLLVLAVFSERSYFALYLCIGWGAPALFIIPWVAVRYTYENTLCWSTNNNMGFWWIIRSSVLLAI
1382    VINFVIFVRIIQILVSKMRAHQMRYTDYKFRLAKSTLTLIPLLGIHEVAFAFLPEETVHGTLRLVKLFFDLFISS
1383    FQGMLVAVLYCFVN
1384
1385    >Clawed frog_ENSXETG00000032385_glp1r
1386    VIYTIGYSLSFSALIIATIILVRFRHLHCTRNYIHLNLFTSFILRAISVFIKDSVLKWMYNLAMNDNQWEGLVSY
1387    QESLSCRLVFAMMQYCVAANYYWLLVEGIYLHTLLVLSVFSEQRLFRLYLCIGWGVPVLFVVPWAIVKYLYEDNG
1388    CWTRNYNMNFWLIIRLPILMAIGRAFVVLVCIMKCMNIDTGECSKSCSSLHDCRLAKSTLTLIPLLGTHEIIFAF
1389    ITDEHAKGALRYIKLFFELSFSSFQGLMVAILYCFIN
1390
1391    >Clawed frog_ENSXETG00000016321_glp2r
1392    IIYTIGYSISLGALLLALVILLLFRKLHCTRNYIHMNLFASFIMRALAVLIKDIVYKNTYFKKNDEMGWMSHLTS
1393    EISTSCRVAQVFMHFFVGANYCWLFVEGLYLHTILVTVILSEKGLLLKYLFIGWFFPLLFVVPWVIAKLYYENNG
1394    CWGVNESPGIWWIIRGPMLLGILINFLIFIKVLKLLYSKLKAQQMRYTDCKYRLARATLALIPLLGMHHVVFTFI
1395    TDELVEGATRHFWLLIQLAFESFQGFVVAIFYCFTN
1396
1397    >Clawed frog_ENSXETG00000018255_gipr
1398    VMYSVGYGISLAALIVAVFILTQLRRLRCTRNLIHCNLFVSFILRGVSLLTRDALLPLHHNMIQGEGDPTNLLRN
1399    RTLVGCRVAQSITQYCVAANYYWLLVEGLYLHNLLVVLSFSEESVLPRYMLLGWGAPVLFVVPWVVVRQLYENSV
1400    CWERNDNYSHWWIIRSPILLAVLINFFIFLRIIRILVLKLRANQMRRSDRKYRLAKSTLTLIPLLGIHEAVFNLL
1401    PEESARGGVRYGKLGAELLLSSFQGLLVAVLYCLCN
1402
1403    >Clawed frog_ENSXETG00000010322_gcrpr
1404    ALYTVGYSVSLLTLISALLILTMCRKLRCTRNSIHANLFASFALRAVSVIVKDVLLAKRWGMQITEVSDWEVLIS
1405    DQAAIGCRIAQVVMQYCILANHYWFVVEAVYLYKLLIGAVFSEKNYYTLYLGWGTPVLFVVPWVTLKYLKENS
1406    ECWALNENMAYWWIIRIPILLASLINLVIFMRILKVILSKLRANQKGYADYKLRLAKATLTLIPLFGIHEVVFIF
1407    ATDEQTSGILRYIKVFFNLFLNSFQGFLVAVLYCFAN
1408
1409    >Coelacanth_ENSLACG00000014981_gcgr
1410    IMYTVGYSLSLAALVLALGILVGFRKLHCMRNYIHINLFVSFILRAVSILVKDALVNTQYKKKIDYENKVQVWLS
1411    GEAMVGCRTAMVLMQYGIAANYYWLLVEGIYLHNLLVIAVFSEKSYFNIYLCIGWGAPVLFVVPWVIVKYLYENI
1412    ECWSKNENMGFWWIIRSPILFAILINFFIFIRIIQILVSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHEVVFVF
1413    ITEEHAQGTLRCIKLFFELFFNSFQGLLVAILYCFVN
1414
1415    >Coelacanth_ENSLACG00000014381_glucagon-like receptor
1416    IIYTVGYSLSIVLLSVAICILLSFRRLHCLRNSIHLNLFVTFLLRAVTVLVKDGLLRQSYSAPYVSPPDWQTFPN
1417    TKALFSCRTAQVLMQYCIGVNYFWLLCEGIYLQALLSASNLSKNNCLRYYILFGWGTPVLFVVPWVTVKYMLENE
1418    ECWLRNISMGVWWIMRAPLLAALIINFLIFIRIFLLLVSRLHSNRLTFSNSKQRLAKVTLTLIPLLGLHETLFAF
1419    VTDESAVGLLRTVKLFYELLLSSVQGCIVTVLYFFTN
1420
1421    >Coelacanth_ENSLACG00000003010_glp2r
1422    FLYTVGYSLSLASLLLAVLILLLMRKLHCTRNYIHINLFCSFILRVIAVFVKDSILDHTYSKRPNNEMGWTSYFK
1423    SQLSMACRATHILMNYFVVANHYWLLVEGIYLHTLLVTVVLSEKRLLQRYILIGWVFPVLFVVPWIITKALYENK
1424    GCWTAQGSFKSGWILWLWKAGESLRIQVNFYIFIKILKLLLSKLKARQLRFSDYKHRLARSTLVLISVFGIQEVV
1425    FAFVTDDQVEGLSRIIRLFIQLPLSSFQYIYIKNIHGFLF
1426
1427    >Coelacanth_ENSLACG00000007720_gipr
```

```
1428    VMYTVGYSLSLAALVLALITLLAFRKLRCTRNYIHMNLFASFILRAISILMRDALLKTHIKQEIKNEGDIFNLLS
1429    DQAAVGCRLAQVLMQYCVGANYYWLLVEGLYLHNLLVVMVLSEKSYFRGYLLIGWGAPVLFVVPWVLVRYFHENT
1430    QCWERNDNMAYWWIIRFPILLAILINFFIFIRIIKILISKLRAHQMRYNNYKCRLAKSTLTLIPLLGIHKLVFEF
1431    VTEEQAKGTLRYVKLFFELFLNSFQGLLVAILYCFVN

1432
1433    >Coelacanth_ENSLACG00000003192_gcrpr
1434    ILYTVGYSLSLLALILALIILGTFRKLHCTRNYIHANLFASFALRAVSVIAKDALLEKRWGMEIMDVTDWGILLS
1435    DEAAIGCRIAQVVMQYCILANHYWFLVEAVYLYKLLIGAVFSEKNYYTLYLYLGWGTPVMFVIPWMAAKYLKENT
1436    ECWGLNENMAYWWIIRFPILMASLINLVIFMRILKVILSKLRANQKGYADYKLRLAKATLTLIPLFGIHEVVFIF
1437    ATDEQTTGTLRYIKVFFTLFLNSFQGFLVAVLYCFAN

1438
1439    >Spotted gar_ENSLOCG00000013910_gcgr
1440    VMYTVGHSLSLGALVLALGILVAFRKLHCMRNYIHMNLFASFILRAVSILVKDALLHPPSSNFTLPKDNEVEKWL
1441    NSTDETPVGCRAALVMMQYSIMANYNWLLVEGIYLHNLLVITVFTERNYFKIYLCIAWGAPLLFVLSWVRLKYLY
1442    ENIQCWERNLNMAVWWVIRSPIQLAILINFFIFIRIIQILVSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHEVV
1443    FAFVTDEHAQGSLRLVKLFFELFSSSFQGLLVAILYCFVN

1444
1445    >Spotted gar_ENSLOCG00000011822_glp2r
1446    IVYTIGYSLSLSSSLSLAVIILLLLRKLHCTRNFIHINLFTSFILRAVVILAKEIILYETYSKRPKDETGWIYILN
1447    SENSPFCRAVQVFMHYLIGANAFWLLVEGIFLHTLLVTPVLSEKRLLKKYMVIGWGTPIMFVVPWAVTKALYENE
1448    GCWRRNTNMGIWWIIRGPIRFSIAVNFYLFIKILKLLLWKLKAEKMTFNDYKFRLARATLVLIPLMGIHEIVFAF
1449    MPDEQIKGRYTRSFIQLTLTSFQGFLVAVLYCFAN

1450
1451    >Spotted gar_ENSLOCG00000014786_gipr
1452    VMYTVGYSLSLAGLTLAFTILLIFRKLRCTRNYIHTNLFASFILRAVSILTRDALLMREAREFGDNRDFVLSDQA
1453    LSGCRVAQVLMQYCVGANYCWLLVEGLYLHNLLVLMVFSENSYFCGYLVIGWGTPVLFVVPWTVVRYLYENKKCW
1454    EMNENMAYWWIIRSPILFSILINFFIFIRIIKILVSKLTAHQMRYTDYKFRLAKSTLTLIPLLGIHEVVFAFITE
1455    EQAAGTLRNVKLFFELFFNSFQGLLVAVLYCFVN

1456
1457    >Spotted gar_ENSLOCG00000003380_gcrpr
1458    VLYTVGYSLSLFTLISALVILLGFRKLHCTRNYIHANLFVSFVLRAVSVIVKDALLEHHWGREITMESDLGEILS
1459    HQAAIGCRIAQVVMQYCILANHYWFFGEAVYLYTVLIGSVFSEKNSCTAYLYLGWGKALP

1460
1461    >Takifugu_ENSTRUG00000012066_gcgra
1462    IMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFILRALSVLIKDALMDTNEKISHWSTTVNNETMI
1463    CCRIAFVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFKIYLCIGWGMPLLFLVPWVMAKYWYENHMCWEL
1464    TTSMNIWWIIRSLILLAVVINFLIFIHIIKILVSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHHMVTIFVTDES
1465    THSTISLRLTKLFIDLFFSSFQGLLVAILYCFVN

1466
1467    >Takifugu_ENSTRUG00000014346_gcgrb
1468    IMYTVGYSLSLGALLLALGILIAFRKLHCMRNNIHMNLFASFILRAVSILVKDAFLTLTLDSRSSNNTQAQAPVN
1469    TTGITWCRGAMVMMQYSVMANNYWLLVEGIYLHSLLVITVFSEKKYFYIYMAIGWGAPLMFVVPWITVKYLYENE
1470    ECWERNINMGFWWIIRSPILFAYLINFFIFIRIIKILMSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTF
1471    VIDESVQKGSLLRLIRLFYDLLFSSFQGLLVAILYCFVN

1472
1473    >Takifugu_ENSTRUG00000017202_glp2r
1474    LISIIGYSLSLFSLTVATLVMAMLRKLHCTRNYIHMNLFVSFILRAMAVILKEIIFYIKHFNLPKDDPGWKSYAD
1475    SAIVLSCRVSAVCMQYFVACNYFWLLVEAVFLHTLLFSAVLTKRRLLKRYMLLGWGTPVLFVTPWTVVKILHENT
1476    GCWSIMNKWIWWIIRGPITLTFVVIFCIFIKILMLLLSKLKADQLKFTDYRYSLVRATLVLIPLLGIHEVVFMVL
1477    TDECMEGRSLYAKNFVNLTLNSFQGFLVAVLYCFAN

1478
1479    >Takifugu_ENSTRUG00000015257_gcrpr
1480    MVYTVGYSISLLTLSTALVILLSFRKLRCSRNYIHANLFLSLILRAVSVIIKDTMLERHWGREIVKQTDVGEMLS
1481    HQAAIGCRMAQAVMQYCVLANHCWFFGEAVYLYSVLIASVFIDNNKHLPYICLGWGTPLLFVIPWVVMKLLKENK
1482    ECWAFNENMNYWWIIRLPILFASLINFLIFMKILKVILSKLRANKIKCLLSYVCRLAKATLTLIPLFGIHEIIFI
1483    FATDEQTTGILRYIKVFFTLTLNSFQGLLVSVLYCYAN

1484
1485    >Tetraodon_ENSTNIG00000010439_gcgra
1486    IMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFIMRALSVLIKDALLDATNKTSHWSTTVNNETMI
1487    CCRIAFVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFKIYLCIGWGMPLLFLVPWVMAKYWYENHMCWEL
```

```
1488    NTSMNIWWIIRSLILLAVVINFLIFIHIIKILVSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHHMVFIFVTDES
1489    TLSTIGLRLTKLFIDLFFSSFQGLLVAILYCFVN
1490
1491    >Tetraodon_ENSTNIG00000012358_gcgrb
1492    VMYTVGYSLSLGALLLALAILVAFRKLHCMRNNIHMNLFASFILRAVSILVKDALLTLTLDSRSSDNTSVHTTVS
1493    ANAAMWCRSAMVLMQYSVIANNYWLLVEGIYLHSLLVITVFSEKKFFHVYMAIGWGAPLMFVTPWIAVKYLFENE
1494    ECWERNINMGFWWIIRSPILFAYLINFFIFIRIIKILMSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTF
1495    VIDESAQKGSLLRTIRLFYDLLFNSFQGLLVAILYCFVN
1496
1497    >Tetraodon_ENSTNIG00000012820_gcrpr
1498    TLYTVGYSVSLLTLSAALIILLSFRKLRCTRNYIHANLFMSLILRAVSVIVKDTMLERHWGREIVKQKDVQEMLS
1499    HQAALGCRIAQAIMQYCVLANHCWFFGEAVYLYSVLIASVFFDNNKHLPYICLGWGTPLLFVTPWVVMKLLKENK
1500    ECWAVNENMNYWWIIRLPILFASLINFLIFIKILKVILSKLRASNQSGYPDFKLRLAKATLTLIPLFGIHEVIFV
1501    FATDEQTTGILRYIKVFFSLTLNSFQGLLVSVLYCYSN
1502
1503    >Tilapia_ENSONIG00000001985_gcgra
1504    IMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFILRALSVLIKDALFETSNIAQGLNRDQEQGFPP
1505    ASIAPVEQLVNNETMISCRIAMVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFKIYQCIGWGTPLIFLVP
1506    WVAIKYLYENQHCWEQNINMKYWWIIRAPILAAVMINFLIFIHIIKILVSKLRAHQMRYTDYKFRLAKSTLTLIP
1507    LLGIHNVVFIFATDESTSGSIGLRLTRLFSDLFFSSFQGLLVAILYCFVN
1508
1509    >Tilapia_ENSONIG00000016459_gcgrb
1510    IMYTVGYSLSLGALLLALGILIFFRKLHCMRNNIHMNLFASFILRAVSILVKDALLTLTLDPRSNSDSQARSWVN
1511    IPAVMWCRGAMVMMQYSVMANNYWLLVEGIYLHSLLVITVFSERKYFYIYLAIGWGAPLIFVLPWIAVKYLYENE
1512    ECWERNINMGYWWIIRSPILFAYLINFFIFIRIIKILMSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTF
1513    VIDESVPKGSVLRLVRLFCDLLFNSFQGLLVAILYCFVN
1514
1515    >Tilapia_ENSONIG00000019565_glp2r
1516    LISVIGYSLSLVSLILATLLMGMLRKLHCTRNYIHMNLFVSFILRAAAILSKEIIMHIMYSNLPKDDPGWNTYSS
1517    SPIVIMCRLSKVCMEYFVACNYFWLLVEAIFLHTLLFTAVLTKRCLLKKYMLLGWGTPALFVTPWTVVKILYENT
1518    ECWSIINRGFWWIIRGPITLSVLVIFFIFIKILMLLLSKLKADQVKFTDYRYSLARATLVLIPLLGIHEVVFTVL
1519    IDECVDGSSRYARNFVNLTLSSFQGFLVAVLYCFAN
1520
1521    >Tilapia_GL831136.1: 8,195 Mb_gcrpr
1522    MLYTVGYSLSLFTLITALIVLLSFRKLHCTRNYIHANLFLSFILRAVAVIVKDTMLEHHWGREIMKPTDVSEMLS
1523    HQAAVGCRIAQVIMQYCVLANHYWFFGEAIYLYSVLIASVFIDSNKYLLYIYLGWGTPLLFVVPWVVVKMLKENK
1524    ECWAVNENMNYWWIIRFPVLLASLINFLIFTKILKVIFSKLRASNPTHYPDYKFRLAKATLTLIPLFGIHEVIFV
1525    FATDEQTTGVLRYIKVFFTLFISSFQGFLVAVLYCFGN
1526
1527    >Tilapia_XP_003459508.2_gipr
1528    VMYTVGYSLSLVSLCVALIILLFFSKLHCTRNYIHSNLFASFILRALSILTKDALLGKTYLEFTDNRDVFEVNSN
1529    QALSSCLVAQVLMHYCVGANYYWLLVEGLYHNLLALMAFSENHFFGGYLLIGWGTPVLFVVPWILVRYMYEDTR
1530    CWEINENMAYWCIIRIPILLAIMVNFFIFIRIILILISKLAHQMRYTDYKFRLAKSTLTLIPLLGIHEVVFAVL
1531    TNVQTDGVFRNINLFFQLFFNSFQGLLVAVLYCFVN
1532
1533    >Sea bass_DLAgn_00191040_gcgra
1534    VMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFILRALSILIKDALLEATNITSQDLGGDQEQGFP
1535    QASMPPVELLVNNETTVSCRIAVVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFKIYLCIGWGTPLIFLV
1536    PWVVAKYLYENQECWEQNINMNYWWIIRSQILLAVVINFLIFIHIIKILVSKLRAHQMRYTDYKFRLAKSTLTLI
1537    PLLGIHQVVFIFVTDESTKTTIGLRLTKLFIDLFFSSFQGLLVAILYCFVN
1538
1539    >Sea bass_DLAgn_00178590_gcgrb
1540    IMYTVGYSLSLGALLLALGILITFRKLHCMRNNIHMNLFASFILRAVSILVKDALLTLTLDPKSSSDSQTQAWVN
1541    IPAVTWCRGAMVMMQYSVMANNYWLLVEGIYLHSLLVITVFSERKYFYIYLTIGWGAPLIFVLPWITVKYLYENE
1542    ECWERNINMGYWWIIRSPILFAYLINFFIFIRIIKILMSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTF
1543    VIDESVPKGSMLRLIRLFCDLLFNSFQGLLVAILYCFVN
1544
1545    >Sea bass_DLAgn_00240050_glp2r
1546    LISVIGYSLSLSSLTLATLLMGLLRKLHCTRNYIHMNLFVSFILRAMAVISKEIILYIMYSNLPKDDPGWNSYSS
1547    SVIALMCKISKVCMEYFVACNYFWLLVEAIFLHTLLFTAVLTKRRLLKRYMLLGWGTPVLFVTPWTVVKILFENT
```

58

```
1548    GCWSIVNRWFWWIIRGPITLSVLVIFFIFIKILMLLLSKLKADQVKFTDYRYSLARATLVLIPLLGIHEVVFTVL
1549    IDECVEGSSRYARNFINLTLSSFQGFLVAVLYCFAN
1550
1551    >Sea bass_DLAgn_00036000_gipr
1552    VMYTVGYCVSLASLSLALIILLFFRKLHCTRNYIHSNLFASFILRAVSILTRDALLSRDTPEINRVLSTVFSNQT
1553    LSGCHVAQVLMQYCVGANYYWLLVEGLYHNLLVVFSDSCYFCGYLLIGWGTPVLFVVPWIIVRYLFENTRCWEI
1554    NENRVYWFIIRTPILLAILINFFIFIRIIHILISKLKAHQMRYTDYKFRLAKSTLTLIPLLGIHEVVFAVLTEEH
1555    TDGVLRNINLFLQLFLNSFQGLLVAILYCFVN
1556
1557    >Sea bass_DLAgn_00100590_gcrpr
1558    MLYTVGYSLSLFTLITALIILLSFRKLHCTRNYIHANLFLSLILRAVSVIIKDTMLERHWGREIMKQTDVREMLS
1559    HQAAIGCRIAQVMMQYCVLANHYWFFGEAIYLYSVLIASVFIDNNKYLPYICLGWGTPLLFVIPWVVMKLLKENK
1560    ECWAVNENMNYWWIIRFPILFASLINFLIFMKILKVILSKLRANNQSGYPDYKLRLAKATLTLIPLFGIHEIIFI
1561    FATDEQTTGVLRYIKVFFTLFLNSFQGFLVSVLYCYAN
1562
1563    >Medaka_ENSORLG00000007082_gcgr
1564    IMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFILRALSILIKDALLEANHTAQDLSRDQDQGFPS
1565    ASMPPMELLVSNMTSVSCRIAVVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFKIYLCIGWGTPLIFLVP
1566    WVILKYLNENQECWEQNISMNYWWIIRAPILLAVVINFLIFIHIIKILVSKLRAHQMRYTDYKFRLAKSTLTLIP
1567    LLGIHQVIFIFVTDESTKGTISLRLTKLFTDLFFSSFQGLLVAILYCFVN
1568
1569    > Medaka_ENSORLG00000013155_glp2r
1570    LISVAGYSLSLFSLSLATLVMGVLRKLHCTRNFIHMNLFVSFILRAVAVMSKEIILHVMYSNLPKDDPGWNTYSS
1571    SPIAVMCKFSKVCLEYFVACNYFWLLVEAIFLHTLLFTAVLTKRRLLKKYMMLGWGTPVLFVTPWTVLKILYENT
1572    GCWLIMNRWFWWIIRGPITFSVLIIFFIFIKILMLLLSKLKADQVKFTDYRYSLARATLVLIPLLGIHEVVFTIL
1573    VDECVEGSSRYARNFINLTLSSFQGFLVAVLYCFAN
1574
1575    >Medaka_ENSORLG00000007568_gcrpr
1576    MLYTVGYSMSLSTLSIALIILLSIRKLHCTRNYIHANLFLSFILRAMAVIIKDTMLDRHWGREIIQQVDVSEMLS
1577    HKAAFGCRAAQVMMQYCVLANHFWFFGEAIYLYSVLISSVLVDKTKYLPYLFLGWGTPLLFVIPWSVMKLLKENK
1578    ECWGANENMNLWWIIRFPILFASLVNFLVFIRILGVIFSKLRASRQRRYPDYKVRLAKATLTLIPLFGIHEVIFL
1579    FVTDEQTTGVLRFTKVFFTLFISSFQGFLVAVLYCFAN
1580
1581    >Amazon molly_ENSPFOG00000002015_gcgra
1582    IMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFILRALSVLIKDALLEANFTSQKISGDRDQGFPS
1583    ESIPPVELLTTISCRTALVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFKIYLCIGWGTPLIFLVPWAVS
1584    KYLYENKECWEQNTDMNIWWIIRAPILLGVVINFIIFIHIIKILVSKLRAHQMRYTDYKVRLAKSTLTLIPLLGI
1585    HQIVFIFLPEETTNKSLHLHLTKLFIDLFFSSFQGLLVAILYCFVN
1586
1587    >Amazon molly_ENSPFOG00000018824_gcgrb
1588    IMYTVGYSLSLGALLLALAILISFRKLHCMRNNIHMNLFASFILRAVSILIKDALLSLMLDPKSGSDAQTQAWVN
1589    IPAVMWCRGAMVMMQYSVIANNYWLLVEGIYLHSLLVITVFSEKKYFIYLAIGWGAPLIFVLPWITVKYLYENL
1590    ECWERNINMGYWWIIRSPILFAYLINFFIFIRIIKILMSKLRAHQMRYTDYKFSRLAKSTLTLIPLLGIHAILFT
1591    FVIDESVPKGSMLRLIRLFCDLLFNSFQGLLVAILYCFVN
1592
1593    >Amazon molly_ENSPFOG00000018904_glp2r
1594    LISVIGYSLSLFSLILATLLMGMLRKLHCTRNYIHMNLFVSFILRAAAVISKEIIFHLMYSNLPKDDPGWNSYSS
1595    SAIVLLCKFSKVCMEYFVACNYFWLLVEAIFLHTLLFTAVLTKRCLLKKYILLGWGTPVLFVTPWTVVKILYENT
1596    GCWSIMNRWFWWIIRGPITLSVLVIFFIFIKILMLLLSKLKADQVKFTDYRYSLARATLVLIPLLGIHEVVFTVL
1597    IDECMEGSSRYARNFINLTLSSFQGFLVAVLYCFAN
1598
1599    >Amazon molly_ENSPFOG00000015868_gcrpr
1600    LYTVGYSLSLSTLIMALIIVLSFRKLHCTRNYIHANLFLSFILRAVAVIVKDSMLDRHLGREIVRHADVTEMLSH
1601    QAAIGCRVAQVMMQYCVLANHYWFFGEAIYLYSVLISCVFIDSNKYLPYLCLGWGTPLLFVVPWAVMKVLKENKE
1602    CWAVNENMNYWWIIRFPVLLASLINFLIFMKILGVIFSKLRANIHSQYPDYKLRLAKATLTLIPLFGVHEVIFIF
1603    ATDEQTTGLLRYIKVFFTLFISSFQGLLVAVLYCYAN
1604
1605    >Stickleback_ENSGACG00000010535_gcgra
1606    VMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFILRALSILIKDALLEANITSQDLSSDQEQGFPQ
1607    ASMPPVELLVNNETAVNCRIAMVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFKIYLCIGWGTPLIFLVP
```

```
1608    WVMAKYLYENQECWGQNINMNYWWIIRSPILLAVVINFLIFIHIIKILVSKLRAHQMRYTDYKFRLAKSTLTLIP
1609    LLGIHQVVFIFVTDESTKATISLRLTKLFSDLFFSSFQGLLVAILYCFVN
1610
1611    >Stickleback_ENSGACG00000018868_gcgrb
1612    IMYTVGYSLSLGALLLALGILISFRKLHCMRNNIHMNLFASFILKAVSILVKDALLSLTLDPRSSADSRTQAWVN
1613    IPAVTWCRGAVVMMQYSVMANNYWLLVEGIYLHSLLVITVFSERKYFFIYLAIGWGAPLMFVLPWITVKYLYENE
1614    ECWERNINMGYWWIIRSPIVFAYLINFFIFIRIIKILMSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTF
1615    VIDESVPKGSMLRLIRLFCDLLFNSFQGLLVAILYCFVN
1616
1617    >Stickleback_ENSGACG0000001479_glp2r
1618    LISVIGYSMSLSSLFIATLLLGMLRKLHCTRNYIHMNLFVSFILRALAVISKEIVLYIMYSKLPKDDPGWNSYSI
1619    SAIALICKFSKVCMEYFVACNFFWLLVEAIFLRTLLFTAVLTKRHLLKNYMLLGWGTPILFVTPWTVVKILYENT
1620    ECWSIVNKWFWWIIRGPITFSVLVIFFIFIKILMLLLSKLKADQLTLTDYRYSLARATLVLIPLLGLHEVVFTVL
1621    IDECVEGSSRYARNFINLTLSSFQGFLVAVLYCFAN
1622
1623    >Stickleback_ENSGACG00000009099_gcrpr
1624    MLYTAGYSLSLFTLVTALIVLLSFRKLHCTRNYIHANLFMSLILRAVSVIVKDTMLERHWGREILKHGDVSEMLS
1625    HQAAIGCRAAQVMMQYCVLANHYWFFGEAIYLYSVLIASVFIDNNKYLPYICVGWGTPLLFVVPWAVMKQLKENK
1626    ECWAVNENMNYWWIIRFPILLVSLINFLIFMKILKVILSKLRASNQSGYPDYKLRLAKATLTLIPLFGVHEVVFI
1627    FVTDEQTTGVLRYIKVFFTLFLNSFQGFLVAVLYCYAN
1628
1629    >Platyfish_ENSXMAG00000012984_gcgra
1630    IMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFILRALSVLIKDALLEANFTSQKISGDRDQGFPS
1631    ESIPPVELLTTVSCRTAVVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFKIYLCIGWGTPLIFLVPWAVS
1632    KYLYENKECWEQNTDMNIWWIIRAPILLGVVINFIIFIHIIKILVSKLRAHQMRYTDYKVRLAKSTLTLIPLLGI
1633    HQIVFIFLPEETTNKSLHLRLTKLFIDLFFSSFQGLLVAILYCFVN
1634
1635    >Platyfish_ENSXMAG00000012523_gcgrb
1636    IMYTVGYSLSLGALLLALAILISFRKLHCMRNNIHMNLFASFILRAVSILIKDALLSPMLDPKSGSDAQTQAWVN
1637    IPAVMWCRGAMVMMQYSVIANNYWLLVEGIYLHSLLVITVFSEKKYFYIYLAIGWGAPLIFVLPWITVKYLYENI
1638    ECWERNINMGYWWIIRSPILFAYLINFFIFIRIIKILMSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTF
1639    VIDESVPKGSMLRLIRLFCDLLFNSFQGLLVAILYCFVN
1640
1641    >Platyfish_ENSXMAG00000005305_glp2r
1642    LISVIGYSLSLFSLTLATLLMGMLRKLHCTRNYIHMNLFVSFILRAAAVISKEIIFHLMYSNLPKNDPGWNSYSS
1643    SAIVLLCKFSKVCMEYFVACNYFWLLVEAIFLHTLLFTAVLTKRCLLKKYILLGWGTPVLFVTPWTVVKILYENT
1644    GCWSIMNRWFWWIIRGPITLSVLVIFFIFIKILMLLLSKLKADQVKFTDYRYSLARATLVLIPLLGIHEVVFTVL
1645    IDECMEGSSRYARNFINLTLSSFQGFLVAVLYCFAN
1646
1647    >Cod_ENSGMOG00000008208_gcgra
1648    IIYTLGYSLSLVALVLALGTLIFFRKLHCMRNNIHMNLFASFILRALSVLIKDALMEDPGLKIRHQDRDFTRDII
1649    TPVDLLVNNETTVACRIAMVMMHFSIMANSYWLLVEGIYLHNLLVIVVFNERNFFNIYLCIGWGAPLLFLLPWVT
1650    AKYLYENEKCWELNVNMNFWWIIRSPILLAVVTNFLIFIHIIKILVSKLRAHQMRYTDYKFRLAKSTLSLIPLLG
1651    IHQMVFVFVLDESSENSLILRFTKVAIDLLVTSFQGLLVAILYCFFN
1652
1653    >Cod_ENSGMOG00000008506_gcgrb
1654    QVVVWCRSAVVMMQYSVMANNYWLLVEGLYLHSLLVTTVFSERNYFYIYLAIGWGAPLAFVLPWATVKYLYENEE
1655    CWERNINMGYWWIIRCPILFAYLINFFIFIRIIQILMSKLKAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTFV
1656    IDESVPKGSMLRLIRLFCDLLFNSFQGLLVAILYCFVN
1657
1658    >Cod_ENSGMOG00000005923_gipr
1659    VVSLVGYSLSLCSLSLATLLMALFRKLHCTRNYIHMNLFGSFILRALSVILKELLLHILYNRTFTDDAGWNSYYN
1660    SATAVMCRVARVAMEYFVACNYFWLLVEAIFLHTLLFTAVLTKRRLLKRYMFLGWVTPVLFVTPWTAIKIQYENR
1661    IIFLIFIKILMLLLSKLKADQGKFTDYRYSLLRATLILFALLGIHEVVFTVLYAESLDGSARYARNFINLTLCSF
1662    QGFVVAVLYCFAN
1663
1664    >Cod_ENSGMOG00000006250_gcrpr
1665    MLYTVGYSLSLSTLTMALLILLSFRKLYCTRNYIHANLFLSFILRAVSVIIKDTLLEHHWGRQIMQQADLHHMLS
1666    HQAAIGCRIAQVMMQYCVLANHFWFFGEAVYLHSVLIGSVFINHNKHLPYICLGWGTPLLFVLPWVLVKLLKENK
1667    ECWALNQNMNYWWIIRLPILLASVINLLIFMKILKVILSKLRASNQNGYPDYKLRLAKATLTLIPLFGVHEVIFV
1668    FATDEQTTGLLRHVKVFFTLFLNSFQGFLVSVLYCYTN
```

```
1669
1670    >Smooth tongue sole_XP_016895711.1_gcgr
1671    IMYTVGYSLSLVALVLALGILIFFRKLHCMRNNIHMNLFASFILRALSILIKDALLVANITSQELSRDQEPGFSR
1672    ASMPPVELLVNNEITVSCRIAVVMMQYSIMANTYWLLVEGIYLHNLLVITVFTERNYFKIYLCIGWGTPLIFLVP
1673    WVIAKYLYENQQCWGQNINMNYWWIIRSPILLAVVINFLIFIHIIKILVSKLRAHQMRYTDYKIRLAKSTLTLIP
1674    LLGIHQVVFIFVTDESTKTTISLRLTKLFIDLFFSSFQGLLVAILYCFVN
1675
1676    >Smooth tongue sole_XP_008313439.1_glp2r
1677    LISIVGYSLSLTSLTLATLLMAVLRKLHCTRNYIHMNLFVSFILRAVAVISKEIVLHVMYTHLPTDDSGWNSYSS
1678    SATAVMCKLSKVCMEYFVACNYFWLLVEAIFLHALLFTTVPTRRRLLKKFMLLGWGTPVIFVTPWMLVKILYENT
1679    ECWSIVNRWFWWIIRGPITLSVLVIFFIFIKILMLLLSKLRADQVKFTDYRYSLARSTLLLIPLLGIHEVVFTVL
1680    IDECVDDSGRYVRNFFNLTLSSFQGFLVAILYCFAN
1681
1682    >Smooth tongue sole_XP_008312440.1_gcrpr
1683    MLYTVGYSVSLFTLVTALIILLSFRKLHCFRNYIHANLFLSFILRAVSVIVKDTMLERHWGREIMKQADVSEMLS
1684    HQAAIGCRIAQVMMQYCVVANHYWFFGEAIYLYSVLIASVFIDNNKYFPYLCLGWGIPLLFVIPWVVMKLLKENK
1685    ECWAVNENMSYWWIIRFPILFASLINFLIFMKILKVILSKLRASQQTGYPDYKLRLAKATLTLIPLFGIHEIIFI
1686    FATDEQMTGFLRYAKVFITLFLNSFQGFMVAVLYCYAN
1687
1688    >Salmon_XP_014038281.1_gcgra
1689    MYTVGYSLSLGALVLALGVLITFRKLHCMRNNIHMNLFSSFILRAVSILIKDVLLDSLSVPLAPGADTHHQTVAG
1690    CRIAMVMMQYSVIANNYWLLVEGIYLHSLLVITVFTERNYFCIYLFIGWGAPLIFVLPWVIVKYLYENEECWERN
1691    INMGYWWIIRSPILLAYLMNFFIFIRIIKILMLKLKAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTLVIDESV
1692    PKGSKMRLIRLFYDLLFNSFQGLLVAILYCFVN
1693
1694    >Salmon_XP_014049938.1_gcgrb
1695    TMYTVGYSLSLGALVLALGVLITFRKLHCMRNNIHMNLFASFILRALSILIKDAMLEAPNVNIGQDQDITHYEVE
1696    WLVNNETAVGCRIAVVMMQYSIMANSYWLLVEGIYLHNLLVITVFTERNYFNIYLCIGWGAPLIFLVPWVMVKYL
1697    YENEECWEQNINMQYWWIIRSPILLAVVINFLIFIHIIKILVSKLRAHQMRYSDYKFRLAKSTLTLIPLLGIHLV
1698    VFVFVTDESTEATIALRLTKLFIDLFFTSFQGLLVAILYCFVN
1699
1700    >Salmon_XP_014035029.1_glp2r
1701    VISIVGYSLSLSSLSLATLLMGILRKLHCTRNYIHMNLFVSFMLRAIAVFIKEIVLHVMYTKLPSDDQGWNSYSN
1702    SVITVVCKASRVSMEYFVACNYFWLLVEAIFLQTLLFTDVLTKRRLLKRYMLIGWGTPFVFVVPWTVSKVLYENK
1703    GCWNNENRWIWWIIRGPITLSVLVIFYIFIKILMLLLSKLKADQVKFTDYRYSLARATIVLIPLLGIHEVVFTIL
1704    IDESVEGSSRYARNFINLTLSSFQGFLVSVLYCFAN
1705
1706    >Salmon_XP_014069579.1_gipr
1707    MYTVGYSLSLASLSLALIILLIFRKLRCTRNYIHTNLFASFILRAISILTRDAVLTRDTPEFRDNRDVSNVLSDK
1708    ALSGCRVAQVLMQYCVGANYYWLLVEGLYLHNLLVLMVFSENSYFCGYLFIGWGSPVLFVVPWIIVRYLYENTRC
1709    WEINENMSYWWIIRTPILLAILVNFFIFIRIIQILVSKLKANQMRYTDYKFRLAKSTLTLIPLLGIHEVVFAVMS
1710    EEQTEGVLRNINLFFELFFNSFQGLLVAILYCFVN
1711
1712    >Zebrafish_ENSDARG00000104022_gcgra
1713    TMYTVGYSLSLGALVLALSILVAFRKLHCMRNNIHMNLFASFILRASSILIKDALSERPDTFPVGQDITTELEVE
1714    WLVKNETAVGCRVAVVMMQYSILANSYWLLVEGIYLHSLLVVTVLTERNYLSIYLSIGWGAPLIFVLPWVIVKYL
1715    YENEECWEQNNHMEYWWIIRSPILLAVLINFFIFIHIIKILVSKLRAHQMRYSDYKFRLAKSTLTLIPLLGIHSV
1716    LFSFVTDESTSHGALPLRLTKLFIDLFFNSFQGLLVAILYCFVN
1717
1718    >Zebrafish_ENSDARG00000036272_gcgrb
1719    AMYTVGYSLSLAALTLALGILVSFRKLHCMRNNIHMNLFGSFMVRALSILIKDTLLDQMNRARSGPSHYQSTRWL
1720    DTQTVVGCRSAMVMMQYSVMANNCWLLVEGLYLHSLLVTTVFSERNYFCIYLCIGWGAPLIFVLPWMTVKYLYEN
1721    EECWERNMNMGYWWIIRSPILFAYLINFIIFIRIIKILMSKLKAHQMRYTDYKFRLAKSTLTLIPLLGIHAVLFT
1722    FVIDESVPKESLLRLIRLFYDLLFSSFQGLLVAILYCFVN
1723
1724    >Zebrafish_XP_009304634.1_glp2r
1725    VLSIVGYSLSFSSLCLAVLIMSLLRKLHCTRNYIHINLFVSFMFRAIAVITKEVILQVAYSNLPRDEVGWNSYTK
1726    SAISFICKASKVSLEYFVGCNYFWLLVEAVFLHTLLFTAVLTRKTLLKKYIFIGWGTPLLFVIPWTVAKTLYENK
1727    SCWMNNIRWIWWIIRGPITLSVIVVIFCIFLKIIRLLLSKLKADQVKFTDYRYSLARATLVLIPLLGVHIVFTLI
1728    IDESVEGSNRYARNFVHLTLSSFQGLIVAVLYCFAN
1729
```

```
1730
1731    >Zebrafish ENSDARG00000025478_gipr
1732    VMYTVGYSLSLASLSLALIILLIFRKLRCTRNYIHTNLFASFILRAVSILTRDALLMKDAPEFRDNKDVSIVLSD
1733    QVMSGCRVAQVLMQYCVGANYYWLLVEGLYLHNLLVLMVFSENSYICVYFFIGWGTPVLFVVPWIIVRYLYENTR
1734    CWEINENMAYWWIIRTPILLAILVNFFIFIRIILILISKLKAHQMRYTDYKFRLAKSTLTLIPLLGIHEVVFAVM
1735    TEEQTEGVLRNVNLFFELFFNSFQGFLVAILYCFVN
1736
1737    >Cavefish_ENSAMXG00000017308_gcgra
1738    TMYTVGYSLSLGALVLALGILVAFRKLHCMRNNIHMNLFASFILRAASILIKDALLERPSLNIDPGITTEMEMEW
1739    LVKNETAIGCRTAVVMMQYSIIANSYWLLVEGIYLHNLLAVTVLTEGNYLSIYLCIGWGAPLIFVLPWIIVKYLY
1740    ENEECWEQNINMEYWWIIRSPILLAVLINFFIFIHIIKILVSKLRAHQMRYNDYKCRLAKSTLTLIPLLGIHLVL
1741    FSLVTDESTSNGAMPLRLTKLFIDLFFNSFQGLLVAILYCFVN
1742
1743    >Cavefish_ENSAMXG00000007858_gcgrb
1744    VMYTVGYSLSLASLILAFSILAAFRKLHCMRNNIHMNLFASFILRAASILIKDALMDTINSAQSGLYTYPELGQN
1745    WVNLEMVTGCRVAVVMMQYSVIANNYWLLVEGLYLHSLLSITVFSERNYFIYLAIGWGAPVIFVLPWVTVKYLY
1746    ENEMCWERNVNMGYWWIIRSPILIAYLINFFIFIRILKILMSKLKAHQMRYTDYKFRLAKSTLTLIPLLGIHAIL
1747    FTFVIDEVSKESLLRLIRLFYDLLFSSFQGLLVAILYCFVN
1748
1749    >Cavefish_ ENSAMXG00000016128_glp2r
1750    IFSVVGYSLSLASLSVAVLIMGLLRKLHCTRNYIHMNLFVSFMFRAMAVITKEIILYTMYSNLPKDENGWNSYSD
1751    STIAVICKASKVFMQYCVGCNYFWLLVEAVFLHTLLFTAVLTKRRLLKRYMLVGWGTPFLFVVPWTVVKILYENK
1752    GCWFINNRWIWWIIRAPITLSVLVIFCIFLKIIKLLLSKLKADQVKFTDYRYSLARATLVLIPLLGIHEIVFTII
1753    VDETIEGHHRYVRNFIHLTLSSFQGFLVAVLYCFAN
1754
1755    >Cavefish_ENSAMXG00000010353_gipr
1756    VMYTVGYSLSLTSLSLALIILLLFRKLRCTRNYIHTNLFASFILRAVSILTRDALLTRDAPEFRDNSDVSSVLSD
1757    QALSGCRIAQVLMQYCVGANYFWLLVEGLYLHNLLVLMVFSENSYFCGYLAIGWGAPVLFVVPWIVMRYLYENTR
1758    CWEINENMAYWWIIRTPILLAILVNFFIFIRIIKILISKLKAHQMRYTDYKFRLAKSTLTLIPLLGIHEVVFAVM
1759    TEEQTEGVLRNVNLFFELFFNSFQGLLVAILYCFVN
1760
1761    >Goldfish_AAS93685.1_gcgra
1762    TMYTIGYSLSLAALVLALAILVAFRKLHCMRNNIHMNLFASFILRASSILIKDAMLERPDEFHVGQDITTELEVE
1763    WLVKNETAIGCRVAMVMMQYSILANNYWLLVEGIYLHSLLVITVLTERNYFAIYQCIGWGAPLIYVLPWVIVKYL
1764    YENEECWVQNINMEYWWIIRSPILLAVLINFFIFIHIIKILVSKLRAHQMRYSDYKLRLAKSTLTLIPLLGIHSV
1765    LFSFVTDESTSHGALSLRLTKLFIDLFFNSFQGLLVAILYCFVN
1766
1767    >Goldfish_AAW82330.1_gcgrb
1768    AMYTVGYSLSLAALTLALAILISFRKLHCMRNNIHMNLFGSFILRAVSILIKDALLDQMNSAQSGPITSRGSAGS
1769    ASDGGQLPAAMVMMQYSVMANNYWLLVEGLYLHSLLVTTVFSERNYFIYLCIGWGAPLIFVLPWMTVKYLYENE
1770    ECWERNINMGYWWIIRSPILFAYLINFIIFIRIIKILMSKLKAHQMRYTDYKFRLAKSTLTLIPLLGIHAILFTF
1771    VIDESVPKESLLRLIRLFYDLLFSSFQGLLVAILYCFVN
1772
1773    >Japanese eel_KI305549.1_gcgra
1774    RKLHCMRNNIHMNLFASFILRALSILIKDALLDTPTLNIRAAAGCRIAMVMMQYSVMANNYWLLVEGIYLHNLLV
1775    VTVFTERNYFNIYLCIGWRCWERNINMGYWWIIRSPILFAVLVSTRLAKSTLTLIPLLGIHVVLFIFVTDDHTPN
1776    GSLRLIKLFYNLFFNSFQ
1777
1778    >Japanese eel_AVPY01196538.1_gcgrb
1779    DRKLHCMRNNIHMNLFASFILLALSILVKDALLDMPSLNIMQDIAPDYEQTAIGCRVAMVLMQYGVMANNYWLLV
1780    EGIYLHNLLVVAVFNERNYFNIYLCIGWGTPLLFLLPWVTVKYLYENEERCWERNINMGYWWIIRCPILLVVLIL
1781    MSKLRAHQMRYTDYKFRLAKSTLTLIPLLGIHGILFAFVIDEQAPSGSLRLTKLFCDLLLNSFQ
1782
1783    >Japanese eel_KI304458.1_glp2r
1784    RKLHCTRNYIHMNLFVSFILRAAAVILKEIIQHIMFTNLPSDETGWNGYPNSMATQVLMHYFVGGNYFWLLVEAI
1785    FLNTLLFTAVLTKRRLLKKYMFIGGTPFLFVVPWTISKILYENQGLAKATLVLIPLLGIHEMVFTFMLDEYIEGQ
1786    TRYIRNFINLTMSSVQGLIVAILYCFAN
1787
1788    >Japanese eel_AVPY01197693.1_gipr
```

```
1789    SQQTAQVRLLRKLRCTRNYIHTNLFISFILRAISILTRDALLRRDTPEFRDDRDVSNVLSDKVKSRILEQHICGV
1790    LSPQALSGCRVALVMQYCVGANYFWLLVEGLYLHNLLVLMVFSENSYFSGYLVVGWGRLAKSTLTLIPLLGIHEV
1791    VFAVMTEEQMEG
1792
1793    >Elephant shark_SINCAMG00000002895_gcgr
1794    VMYTVGYSLSLGTLILALGILVGFRKLHCMRNYIHMNLFASFILRAVSILIRDALFKMHYTSVTTGNPDIKMWFN
1795    NKTAGGCKAAQVLMQYCIGANYYWLLVEGVYLHNLLVIAVFSEKSFFNIYLYIGWGAPVLFVVPWVVVKYLYENS
1796    GCWTLNENMAFWWIIRCPILLSILINFFIFIRIIHILVSKLRAHQMRYNDYKFRLAKSTLTLIPLLGIHEVVFAF
1797    FMDEHAQGTLRLVKLFFDLFLNSFQGMLVAILYCFVN
1798
1799    >Elephant shark_SINCAMG00000015132_glp1r-like
1800    FVYTAGYTLSLISLVAANLTLISLRKLRCTRNYIHVNLFSSFILRAISVLLKDSFTDSPDSESVPTEGTHSTGSD
1801    TVPCRAAQLLLHYSVASSWYWLLVEGIFLYTLLVLSVFSPTKSYRLYLMIGWGLPLAPLIPWTLVKYFQENSGCW
1802    RQNSNMGFWWILRAPLLLVICINVAIFVRIMMLLISKMRSHQMQNTDFRYRLTKSTLILIPMLGVHEIVLALVTD
1803    EMAEGTLRYLKLSVELFFGSVQGFLVAVLYCFIN
1804
1805    >Elephant shark_SINCAMG00000004406_glp2r
1806    LIYTVGYSFSLCSLSLAMAILLWLRKLHCTRNYIHMNLFVSFMLRALAVLIKDILLHSTYSKRPDDESGWISFFN
1807    VEAKCLFQISNGCKVAQVFMHYFVGANFFWLLVEGIYLHKLIVLAVLSEKNLLKQYILIGWVCPVLFVVPWIITK
1808    ITSENEGCWGRNVNMWIWWIIRGPVSIAIIMNFYFFLKIMKMLLSKLQAQQMRFGDYKYRLARSTLVLIPLLGIH
1809    ELVFIFLMDEHVEGLLRHIRLFLQLVISSFQGFLVAVLYCFTN
1810
1811    >Japanese lamprey_JL7967_glp1r
1812    PAFCEGCKGGGRLSLSLCVVCDHRRLHCTRNYIHLNLFLSFILRAISILVKDAVLRSIYASDSDTGAWDEVLAHQ
1813    SAACKAVFVLMQYCVAANYYWLLVEGVYLHTLLALAVFSERRSFKAYVAIGWGERPQLRAEGGGSFRWLHEASGG
1814    SRFPRKLQAIQKDPDILLTCMCSRHVHETP
1815
1816    >Japanese lamprey_JL5602_glp2r
1817    LVYTTGYSVSLASLVVAMGILIMFRRLHCTRNYIHMNLFASFILRAISVLLKDAALTHNYTKRPQNASDWLGYFG
1818    YEGYTGCRATHVFMQYCIGANYFWLLVEGSYLHTLLTMCVFTEKKLIAFYVLLGWGTPIMFIIPWMVSKLIYENQ
1819    GCWGINTWMGIWWIIRGPILFSIVVNFILFVKIIKILLSKLKAQQSKFSNYKSRLARSTLTLIPLLGIHEVVFIF
1820    LADEHATGRTRYVRLFIQLSFSSIQGFFVALLYCFSN
1821
1822    >Japanese lamprey_JL10168_gcrpri
1823    SLACVRVRRVQAAAGCRTAQVLTQYCIVANYYWLLVEGVYLHTLLVVPVFSEETYFKVYLAVGWGTPVLFVAPWV
1824    VVKYLKENTECWGHNENMGYWWIIRTPIHLAILINFIMFIRILKIITSKLRAHQMGYTDFKFRSRLAKSTLTLIP
1825    LLGIHEIVFTFITDEHAHDALRHVKLFFALFFNSFQVRPLALPFYTGI
1826
1827    >Japanese lamprey_JL9019_gcrprii
1828    KMYTVGYSLSLAALVLALATFLMFRRLHCTRNYIHMNLFGSFILRAMSILVKDALLETHWEVDPSNATEWGAFLS
1829    AEAAVSCRTAQVLMQYCIVANYCWLLVEGLYLHTLLSVTVFSEEAHFRLYLLIGWGEHRASSSSSSSSSSSCIIYY
1830    MYNVQ
1831
1832    >Sea lamprey_ENSPMAG00000003877_glp2r_GL484465
1833    YTTGYSVSLASLVVAMGILIMFRRLHCTRNYIHMNLFASFILRAISVLLKDAALTHNYTKRPQNASDWLGYFGYE
1834    GYTGCRATHVFMQYCIGANYFWLLVEGSYLHTLLTMCVFTKKKLIAFYVLLGWGTPIMFIIPWMVSKLIYENQGC
1835    WGINTWMGIWWLIRGPILFSIVVNCILFVKIIKILLSKLKAQQSKFSNYKSR
1836
1837    >Sea lamprey_ENSPMAG00000008188_gcrpri_GL476704
1838    QAAAGCRTAQVLTQYCIVANYYWLLVEGVYLHTLLVVPVFSEETYFKVYLAVGWGTPVLFVAPWVVVKYLKENTE
1839    CWGHNENMGYWWIIRTPIHLAILINFIMFIRILKIITSKLRAHQMGYTDFKFRSRLAKSTLTLIPLLGIHEIVFT
1840    FITDEHAHDALRHVKLFFALFFNSFQVRPLA
1841
1842    >Sea lamprey_ENSPMAG00000003562_gcrprii_GL480239
1843    KMYTVGYSLSLAALVLALATFLMFRRLHCTRNYIHMNLFGSFILRAMSILVKDALLET
1844
1845    >Ciona_ENSCING00000006559_gcgr-like
1846    KIYTAGYIFSLVCMVIALFILMFFKKLHCTRNYIHMNLMLSFIVRYVAVMVKDKVLEDHYAVGQTNLTQMEMSQY
1847    CDDVAGTDGLMVSCRLVITLMHYAIIANYFWLLVEGVYLQLLLVFVMTEYKYFPIFMAFGWGAPWIPIGIWVAFR
1848    ITFENVGCWEIYNALRTWWILSAPILISIAINFIIFINIIRMIVSKLRANNMTRSDYKYRLARSTLALIPLLGIH
1849    YIIFMGVSDSVTNSTITKTKFAFELILTSLQGSIIAILYCFLN
```

1850

1851   >Ciona_ENSCING00000006557_gcgr-like

1852   DIYTAGYIFSLVCMVIALFILMFFKKLHCTRNYIHMNLMLSFIVRYVAVMVKDKVIEDHYAVGRPNLTQMEMSQY

1853   CDDVAGTDGLMVSCRLVITLMHYATIANYFWLLAEGVYLQLLLVFVMTEYKYFPIFMAFGWGATWIPIGIWIAFR

1854   ITFENVGCWEVNNMIPIWWILRAPILISIAINFIIFINIIRMIVSKLSANNMTRSDYKYRLARSTLALIPLLGIH

1855   YIVFMGVSDSVTDNSAFINTKFAFEIILTSLQGSIIAILYCFLN

1856

1857

1858