Biochimica et Biophysica Acta (BBA) – General Subjects
March 2000; 1474(1) : 61-69
http://dx.doi.org/10.1016/S0304-4165(99)00213-5

Molecular cloning, tissue distribution and sequence analysis of complete glucokinase cDNAs from gilthead seabream (Sparus aurata), rainbow trout (Oncorhynchus mykiss) and common carp (Cyprinus carpio)[1]

S. Panserat[a,], C. Blin[b], F. Médale[a], E. Plagnes-Juan[a], J. Brèque[a], J. Krishnamoorthy[b] and S. Kaushik[a]

[a] Laboratory of Fish Nutrition, INRA-IFREMER, 64310 St-Pée-sur-Nivelle, France
[b] INSERM U458, Hôpital Robert Debré, 75019 Paris, France

*panserat@st-pee.inra.fr Fax: +33-5-5951-5452

Abstract: The enzyme glucokinase (GK) (EC 2.7.1.1) plays an important role in the control of glucose homeostasis. Qualitative and/or quantitative variations in GK enzyme have been postulated by previous studies to explain why dietary carbohydrate utilisation is lower in gilthead seabream (Sparus aurata) and rainbow trout (Oncorhynchus mykiss) than in common carp (Cyprinus carpio). In this study, we report the isolation and characterisation of a full-length cDNA coding for GK in these teleosts. Amino acid sequences derived from these cDNA clones are highly similar to other vertebrate GKs. These findings, including a detailed phylogenetic analysis, reveal that GK gene highly homologous to mammalian GK exists in these fish species with similar tissue specific expression (mainly liver).

Keywords: Fish nutrition; Dietary carbohydrate; Glucose phosphorylation

INTRODUCTION

In vertebrates, glucokinase or hexokinase IV (GK, E.C. 2.7.1.1) from hepatic and pancreatic tissues plays a important role in controlling the rate of glucose utilisation in both cell types as well as in glucose homeostasis [1, 2, 3, 4, 5, 6, 7]. In fish, it seems that there is no strict control of blood glucose level [8, 9], with both oral administration of glucose as well as ingestion of high carbohydrate diet resulting in hyperglycemia [10, 11]. Given that this can not be due to insulin deficiency in fish [12], one of the currently admitted hypotheses to explain the low dietary carbohydrate utilisation in fish is its inability to convert efficiently the intracellular glucose into glucose-6-phosphate due to the absence of an inducible hepatic GK activity [8, 9]. Literature data in fish on the existence of a functional GK-like enzyme and on the induction of GK expression by dietary carbohydrates have been rather contradictory [13, 14, 15, 16, 17, 18]. We recently reported isolation and characterisation of hepatic cDNA sequences from three cultured teleosts which are homologous to a portion of mammalian GK sequences [19]. These three teleosts differ in their capacity to utilise dietary carbohydrates : an omnivorous fish, namely common carp, able to use efficiently high levels of dietary carbohydrates and two carnivorous species, rainbow trout and gilthead seabream, less tolerant to dietary carbohydrates [8, 9]. In the present work, we obtain the full-length cDNA sequences of GK from all these three teleosts and addressed the issues of tissue-specific GK gene expression in these species. Furthermore, by phylogenetic analysis, we show that the fish GKs are closely related to GKs from vertebrates and distinct from other hexokinases including the « bonafide » GK from yeast (E.C. 2.7.1.2).

MATERIAL AND METHODS

Fish, diets and RNA isolation

Tissue samples (liver, muscle, heart, kidney, brain) were obtained from rainbow trout *(Oncorhynchus mykiss)* at the INRA experimental fish farm (Donzacq, France), from common carp *(Cyprinus carpio)* and gilthead seabream *(Sparus aurata)* at the ICBAS experimental fish farms (Vila Real and Olhao, Portugal). Juvenile immature fish (body weight range at the end of the growth period : about 150 g) were grown for 10 weeks at 18°C during spring under natural photoperiods. They were fed twice a day to near satiation with formulated dry diets containing high levels of digestible carbohydrates (>20%). On the day of sampling, fish were fed once and sacrificed 6 hours after feeding. Tissues were clamp frozen (nitrogen liquid) and stored at -80°C. Total RNA was extracted from common carp, rainbow trout and gilthead seabream livers and other tissues as described by Chomczinski and Sacchi [20]. PolyA mRNAs were purified from total RNAs using a poly dT column according to the manufacturer advice (Promega, USA).

Reverse transcription (RT), Rapid amplification of the cDNA extremities-polymerase chain reaction (RACE-PCR) and molecular cloning of PCR fragments

The 5’ and 3’ cDNA extremities were determined by the RACE-PCR method as detailed in the manufacturer’s notice (Boerhinger, Roche Molecular Biochemicals, Germany). The teleost GK specific primers were designed from the partial sequence data of the same species previously obtained in our laboratory [19] (Table 1). Using the reverse transcription system, cDNA was synthesized by incubating 1µg of polyA mRNA from fish livers with

AMV reverse transcriptase for 1h at 42°C using either the oligodT primer (3' Race) or a species-specific GK primer (5' Race) (Table 1). The RACE-PCR reactions were carried out at 59°C of annealing temperature using species-specific primers (Table 1). PCR products were subjected to electrophoresis in 1% agarose gels, hybridized with labeled GK probes and the relevant fragments were purified (Micropure System, Amicon, USA). These purified DNA fragments were inserted into the pCR^{TM}II-TOPO plasmid (Invitrogen, USA) and used for transformation of One Shot^{TM} competent cells (Invitrogen, USA). Clones with inserts were selected by EcoRI digestion of the plasmid DNA and were sequenced using the dideoxynucleotide chain termination method [21] (Sequenase-2 sequencing kit, Amersham, UK).

Northern and RT-PCR analysis.

Samples of 20 μg of total RNA samples were submitted to electrophoreses on 1% agarose gels containing 5% formaldehyde and capillary transferred onto nylon membrane (Hybond-N^{+}, Amersham, UK). Membranes were hybridized with [^{32}P] DNA specific for GK sequences labeled by random priming (Stratagene, USA) [19] (Genbank accession numbers AF053330, AF053331 and AF053332 for gilthead seabream, rainbow trout and common carp GK related probes respectively). Membranes were also hybridized with a carp 16 S ribosomal RNA probe (Genbank accession number MICCCG) to check the inter-sample variation in loading. After stringent washing, the membranes were exposed to X-ray film and signal intensity autoradiograms were assessed using Visio-Mic II software (Genomic, France).

By annealing 2µg of total RNA with 1µg of random primers and incubating with AMV reverse transcriptase (Invitrogen, USA) for 1h at 42°C, total cDNA was synthesized. Using specific primers derived from the previously acquired partial GK cDNA sequences [19] (Table1), GK-specific cDNA fragments were synthesized. At this end, a 35 cycle PCR reaction was carried out in a final volume of 25 µl containing 1.5 mM $MgCl_2$, 4 pmol of each primer, 2 µl total cDNA and 1 U of Taq polymerase (Boerhinger, Roche Molecular Biochemicals, Germany) with an annealing temperature of 51°C for common carp and gilthead seabream, or 55°C for rainbow trout.

Sequence analysis

Nucleotide sequences were compared with those from the Genbank database with the basic local alignment search tool (BLAST) algorithm [22]. Amino acid sequence alignments were assessed with the Clustal-W multiple alignment algorithm [23]. Percentage of amino acid conservation between GKs was performed by Align program [24]. Amino acid alignments of HK from various eucaryotes (Table 2) were used to construct a phylogenetic tree with the PAUP (phylogenetic analysis using parsimony) algorithm [25]. An heuristic search was performed with the TBR (tree bisection reconnection) branch-swapping algorithm. A tree was produced representing 50% consensus of 1000 replicates. Of the 997 positions, 770 were variable and 623 were phylogenetically informative.

RESULTS

In order to obtain a full-length cDNA sequence data for GK from the three teleosts, we used an established strategy called Rapid Amplification of cDNA Extremities (RACE-PCR). This was possible because we already had the partial sequence of these cDNAs [19]. Precisely, the strategy involved obtaining of full-length sequence information from two substantially overlapping 5' and 3' fragments of a given cDNA (Figure 1). Analysis of these sequences revealed the following features:

First, the initiator codon, arbitrarily fixed on the first ATG, resulted in an open reading frame of 478, 471 and 476 amino acids for gilthead seabream, rainbow trout and common carp respectively (Figure 1). These amino acid and nucleotide sequences were compared to sequence data bases using the BLAST algorithm and correspond unequivocally to GK sequences ($p=10^{-70}$ to 10^{-86}) except the 5' and 3' untranslated regions. Deduced teleost amino acid GK sequences were aligned with the human liver GK sequence using the Clustal-W algorithm (Figure 2). The teleost GK sequences share about 88% amino acid identity and bear high similarity with GK sequences from higher vertebrates (up to 80%) (Table 3). Such strong similarity is also noted for the amino acid residues critical for enzyme activity (Figure 2).

Secondly, although the derived GK cDNA size of 2070 bp and 2655 bp respectively for gilthead seabream and rainbow trout (Figure 1) are in agreement with that deduced from Northern blot analysis of mRNA from these species (Figure 3a), it was not possible to

assess the mRNA size for the common carp as no detectable GK mRNA was found in Northern blot (Figure 3a). Indeed, we found a low level of GK gene expression associated with a low GK activity in common carp even when fed diets rich in carbohydrates [26].

Thirdly, from rainbow trout mRNA, two distinct 3' segments were obtained by RACE-PCR of about 1.8 and 2.2 kb in size. Sequences of these two fragments were identical excepting an additional sequence of 348 bp at the 3' end of the longer fragment. This suggests that rainbow trout has two distinct cDNAs that differ by their 3' untranslated region probably due to the use of less consensual polyadenylation recognition signal (AGTAAA) as noted in the Figure 1c. In fact, the Northern blot data confirmed the presence of two GK mRNA species of 2.4 and 2.7kb in rainbow trout liver (Figure 3b), the latter being the major form. This confirms that the optimal AATAAA sequence is the major polyadenylation recognition signal.

Fourthly, GK mRNA expression studied by RT-PCR (more sensitive than Northern blot) in different tissues (liver, muscle, heart, brain and kidney for common carp and rainbow trout; only liver and muscle for gilthead seabream) revealed that GK mRNA expression is highly specific to liver (for all three species) and brain (for the rainbow trout) (Figure 4).

Finally, the evolutionary relationship among the GKs was investigated by the construction of a phylogenetic tree (Table 2 and Figure 5). Alignment of 25 eucaryotic GK amino acid sequences (including the three teleost sequences) was performed with the clustal-w algorithm [23] and the PAUP algorithm [25] was used to produce an unrooted tree. This

tree has 2632 steps with a consistency index of 0.854 and after excluding uninformative characters, an index of 0.837. Hexokinases from unicellular eucaryotes or from plants are clearly separated from that of multicellular eucaryotes (bootstrap value of 97%) and in this last group, vertebrate HKs clustered together, with a 98% bootstrap value. The yeast GK considered as the « bonafide » GK given its unique substrate specificity for glucose is divergent from all other sequences. Within the eucaryotic group, vertebrate HKs are clustered according to the enzyme type (I to IV) (bootstrap value 99 and 100%) and all the teleost amino acid sequences are related to the type IV group, with a 100% bootstrap value.

DISCUSSSION

« Bonafide » glucokinase enzyme as defined by its unique substrate (glucose) specificity has been found only in the yeast [27]. However, in vertebrates, type IV hexokinase is designated as glucokinase (GK) which is responsible for postprandial regulation of glucose homeostasis [28, 29]. Presence of this enzyme in fish has long been suggested since remained controversial [13, 14, 15, 16, 17, 18]. Our study establishes unequivocally the existence in teleosts of DNA and mRNA sequences homologous to the vertebrate type IV enzyme (GK) further confirmed by cluster analysis in phylogenetic studies. In mammals, GK gene expression is restricted to liver, pancreas and some neuroendocrine cells of the brain although the role of GK in this latter tissue is unknown [30]. Our data demonstrate that, in teleosts, GK gene expression is highly specific to liver. Interestingly, also in trout brain, the GK gene expression was found. Although direct evidence that these teleostean GK cDNA correspond to functional GK enzyme is lacking, the nucleotide and amino acid sequence homology with mammalian GK sequences [31], conservation of critical amino acids involved in glucose and ATP binding [31] and own observation of a high hepatic GK activity in teleosts fed with carbohydrates [26] are in favour of the existence of functional GK enzymes in these species.

The GK region located between the ATP and glucose binding sites (Figure 2) are not totally similar between the full-length GK cDNAs and previously characterised partial GK cDNAs [19] : while complete nucleotide identity is observed for gilthead seabream sequences, only 99% and 96% nucleotide similarities are noted for common carp and

rainbow trout respectively. This observation suggests existence of polymorphisms in fish GK gene (RNA preparations used for the initial cloning [19] and the present Race-PCR were from different animals). Detailed screening of teleost GK gene polymorphisms may provide insight into the possible existence of different forms of GK enzymes possessing distinct catalytic properties as it has been observed in mammals [1, 2]. Besides the qualitative analysis of the GK enzyme in fish, the present data do not exclude the possibility that other minor GK mRNA species may be present. Indeed, in this study, for rainbow trout, two mRNA species are present together only when the total GK expression is high (data not shown). It may suggest that under conditions of high level of GK gene transcription utilisation of cryptic polyadenylation signal (AGUAAA) may become prominent. Naturally occurring variants of consensus AAUAAA sequence in humans indicate that changes in the second nucleotide position are relatively well tolerated (as reported here) with respect to signal function whereas mutations in any other position inhibit the RNA processing [32, 33]. Although majority of eucaryotic gene transcription units possess a single polyadenylation signal, numerous examples of transcription units with multiple poly(A) signals, all within a single 3'-terminal exon, have been described over the past several years [32, 33]. The physiological significance of two mRNA species in rainbow trout remains unknown. However, if different forms of mRNAs have different stability or translation efficiency, then the use of alternative poly(A) sites can have positive or negative impact on the protein expression.

It can be surprising that hepatic GK enzyme, an enzyme involved in glucose homeostasis in vertebrates [1], was so conserved in phylogenetically far distinct (carnivorous) animals

such as gilthead seabream and rainbow trout. We hypothesize that utilization of dietary carbohydrates as a source of energy by fish can be important during fish development specially at the embryo or larval ages: indeed, young larvae of seabass *D. Labrax* exhibit high specific activities for amylase, an enzyme involved in carbohydrate digestion, but this activity declines during development [34]. In this context, the presence of a functional GK enzyme in these species could be vital during early ontogenesis of fish and can explain its genetic conservation. Further studies are necessary to describe the ontogenesis of GK gene expression in rainbow trout and gilthead seabream. In contrast, the existence of functional GK in omnivorous common carp can be due to its « natural » carbohydrate-rich feeding. In this case, a functional GK enzyme involved in dietary glucose utilization is necessary. However, the low induction of GK gene expression by dietary carbohydrates (observed also for the GK activity [26]) is probably linked to an inherent strict control of glycemia as generally observed in omnivorous fish [9].

The phylogenetic analysis of eucaryotic HKs confirms the existence of a type IV HK gene before the separation of the vertebrates into marine and terrestrial animals, about 350 million years ago. Our results also show that the type IV group got separated from the others before the event of duplication and fusion that led to HKI to III, as proposed by Cardenas et al [27]. Finally, the phylogenetic relationship between the three teleost GKs confirm the existence of at least two well defined branches of teleostei as has been suggested by the previous analysis with the partial GK cDNA sequences as well as by other studies [19, 35], defining the ostariophysi super-order (including cypriniform order (common carp)) and the neognathi super-order (including salmoniform (rainbow trout) and

perciform (gilthead seabream) orders). Overall, relationship among these three teleosts observed in this study is in agreement with the notions derived from classical morphometric [35] and genetic analyses.

Altogether we demonstrate the presence of type IV HK (GK) gene expression in teleosts and the data obtained can be put to use for improving the efficiency of dietary carbohydrate utilisation by fish [9]. However, as shown by the present data and our previous study [26], poor dietary carbohydrate utilization in rainbow trout probably involves other protein(s) either in liver or in other tissues than GK alone. Indeed, Glut4 glucose transporter was recently reported to be absent in muscle of tilapia [36] and there is also generally a low number of insulin receptors in the muscle of rainbow trout [37]. Globally, the exact contribution of liver in comparison with peripheral insulin-sensitive tissues (skeletal muscle and adipose tissue) to the observed hyperglycemia in « carnivorous » fish requires further studies.

Acknowledgements:

We are grateful to Dr F. L'Horset for helpful discussions and encouragement. We thank F Vallée, F Terrier, P. Rema and J. Santinha for the maintenance of the rainbow trout (Inra experimental facilities), common carp (Vila Real, Portugal) and gilthead seabream (Olhao, Portugal). This work was supported by the Aquitaine Region (N°CCRRDT: 960308003) and the European Commission (Fisheries Agricultural and Agro-Industrial Research, Contract FAIR N°CT95-0174).

References

[1] J.E.Wilson, Rev. Physiol. Biochem. Pharmacol. 126 (1995) 65-198.

[2] N. Vionnet, M. Stoffel, J. Takeda, K. Yasuda, G.I. Bell, H. Zouali, S. Lesage, G. Velho, F. Iris, P. Passa, P. Froguel, D. Cohen, Nature 356 (1992) 721-722.

[3] A. Gruppe, B. Hultgren, Y.H. Ryan, M. Bauer, T.A. Stewart, Cell 83 (1995) 69-78.

[4] K.D. Niswender, M. Shiota, C. Postic, A.D. Cherrington, M.A. Magnuson, J. Biol. Chem. 272 (1997) 22570-22575.

[5] F.M. Matchinsky, B. Glaser, M.A. Magnuson, Diabetes 47(3) (1998) 307-315.

[6] C. Postic, M. Shiota, K.D. Niswender, T.L. Jettob, Y. Chen, J.M. Moates, K.D. Shelton, J. Lindner, A.D. Cherrington, M.A. Magnuson, J. Biol. Chem. 274 (1) (1999) 305-315.

[7] D.W. Piston, S.M. Knobel, C. Postic, K.D. Shelton, M.A. Magnuson, J. Biol. Chem. 274 (2) (1999) 1000-1004.

[8] CB. Cowey, M Walton, in: J.E. Halver (Ed.), Fish Nutrition, Intermediary metabolism, Academic Press, New-York, (1989) 259-329.

[9] R. Wilson, Aquaculture 124 (1994) 67-80.

[10] T.N. Palmer, B.E. Ryman, J. Fish Biol. 4 (1972) 311-319.

[11] F Bergot, Comp. Biochem. Physiol. 64A (1979) 543-547.

[12] T.P. Mommsen, E.M. Plisetskaya, Rev. Aquat. Sci. 4 (1991) 225-259.

[13] F. Nagayama, H. Ohshima, Bull. Jap. Soc. Sci. Fish. 40 (1974) 285-290.

[14] C.B. Cowey, D. Knox, M. Walton, J. Adron, Br. J. Nutr. 38 (1977) 463-470.

[15] F. Nagayama, H. Ohshima, H. Suzuki, and T. Ohshima, Biochem. Biophys. Acta 615 (1980) 85-93.

[16] M. Furuichi, Y. Yone, Bull. Jpn. Soc. Sci. Fish 48 (1982) 463-466.

[17] M.A. Tranulis, B. Christophersen, B. Borrebaek, B. Comp. Biochem. Physiol. 116B (1997) 367-370.

[18] M.A. Tranulis, O. Dragni, B. Christophersen, A. Krogdahl, B. Borrebaek, Comp. Biochem. Physiol. 114B (1996) 35-39.

[19] C. Blin, S. Panserat, F. Médale, E. Gomes, J. Breque, S. Kaushik, R. Krishnamoorthy, Fish Physiol. Biochem. 21(2) (1999) 93-102.

[20] P. Chomczinski, M. Sacchi, Analytical Biochem. 162 (1987) 156-159.

[21] F. Sanger, S. Nicklen, A.R. Coulson, Proc. Natl. Acad. Sci. USA 74 (1977) 5463-5467.

[22] S. Atlschul, W. Gish, W. Miller, E. Myers, D. Lipman, J. Mol. Biol. 215 (1990) 403-410.

[23] D. Higgins, P. Sharp, C.A.B.I.O.S. 5 (1989) 151-153.

[24] C. Myers, E. Miller, C.A.B.I.O.S. 4 (1989) 11-12.

[25] D.L. Swofford, (1990). PAUP: Phylogenetic Analysis Using Parsimony, version 3.0 Illinois Natural History Survey, Champaign.

[26] S. Panserat, F., Medale, C. Blin, J. Breque, C. Vachot, E. Plagnes-Juan, E. Gomes, R. Krishnamoorthy, S. Kaushik, Am. J. Physiol. (1999) (submitted).

[27] M.L. Cardenas, A. Cornish-Bodwen, T. Ureta, Biochim. Biophys. Acta 1401 (1998) 242-264.

[28] M.A. Magnuson, J. Cell. Biochem. 48 (1992) 115-21.

[29] R.L. Printz, M.A. Magnuson, D.K. Granner, Annu. Rev. Nutr .13 (1993) 463-96.

[30] T.L. Jetton, Y. Liang, C.C. Pettepher, E.C. Zimmerman, F.G. Cox, K. Horvath, F.M. Matshinsky, M.A. Magnuson, M.A. J. Biol. Chem. 269 (5) (1994) 3641-3654.

[31] M. Veiga-Da-Cunha, S. Courtois, A. Michel, E. Gosselain, E. Van Schaftingen, J. Biol. Chem. 271 (1996) 6292-6297.

[32] G. Edwalds-Gilbert, K.L. Veraldi, C. Milcarek, Nucl. Acid. Res. 25(13) (1997) 2547-2561.

[33] D.F. Colgan, J.L. Manley, Gene Dev. 11 (21) (1997) 2755-2766.

[34] C.L. Cahu, J.L. Zambonino Infante, Comp. Biochem. Physiol. 109A (1994) 213-222.

[35] C.R. Gilbert, in : Physiology of Fishes, Evolution and Phylogeny, CRC Press Inc. (1993) 1-34.

[36] J.R. Wright J.R., W. O'Hali, H. Yuang, X. Han and A. Bonen, Gen. Comp. Endocrinol. 111 (1998) 20-27.

[37] M. Parrizas, J. Planas E.M. Plisetskaya and J. Gutierrez, Am. J. Physiol. 266 (1994) R1944-R1950.

Legends

Figure 1 : Nucleotide sequences of the three teleost GK cDNA and deduced amino acid sequences : **a** : gilthead seabream, **b** : common carp, **c** : rainbow trout. Underlined letters correspond to the poly(A) cleavage signal. For rainbow trout, the less consensual poly(A) addition site has been also indicated in italic followed by the beginning of the poly(A) tail (#). Bold letters indicate the first amino acid and the codon stop.

Figure 2 : Comparison of the three teleost and human amino acid sequences deduced from GK cDNA sequences. The stars are the conserved amino acid residues between the four sequences. In human sequence, bold letters correspond to the amino acids that bind glucose. Boxes correspond to a part of the ATP-binding site and to a position of the glucose-binding site from where the degenerated primers were designed to characterise the partial GK cDNA clones in a previous study [19].

Figure 3 : GK gene expression in fish fed with carbohydrates (at 6h after the meal). Northern blot analysis : **a**) comparison of the GK mRNA sizes between rainbow trout and gilthead seabream. The 16S probe served as an internal control of sample loading. No GK mRNA was detected in common carp. 1-2 : rainbow trout samples ; 3-4 : gilthead seabream samples ; 5-6 : common carp samples. **b**) existence of two mRNA GK species for different rainbow trout samples.

Figure 4 : Tissue specificity of GK gene expression in fish fed with carbohydrates (at 6h after the meal) by RT-PCR analysis. L : Liver, M : (white) Muscle, K : kidney, H : Heart, B : Brain. - : negative control.

Figure 5 : Phylogenetic analysis of the teleost GK amino acids sequences deduced from cDNA sequences. Accession numbers of all the GK cDNA sequences are given in Table 2. Sequences were aligned using the Clustal-W algorithm and were analysed by phylogenetic analysis using the parsimony (PAUP) algorithm. The evolutionary relatedness of the HKs is proportional to the length of the horizontal bars. The numbers on the branchs indicate the bootstrapping value for the node.

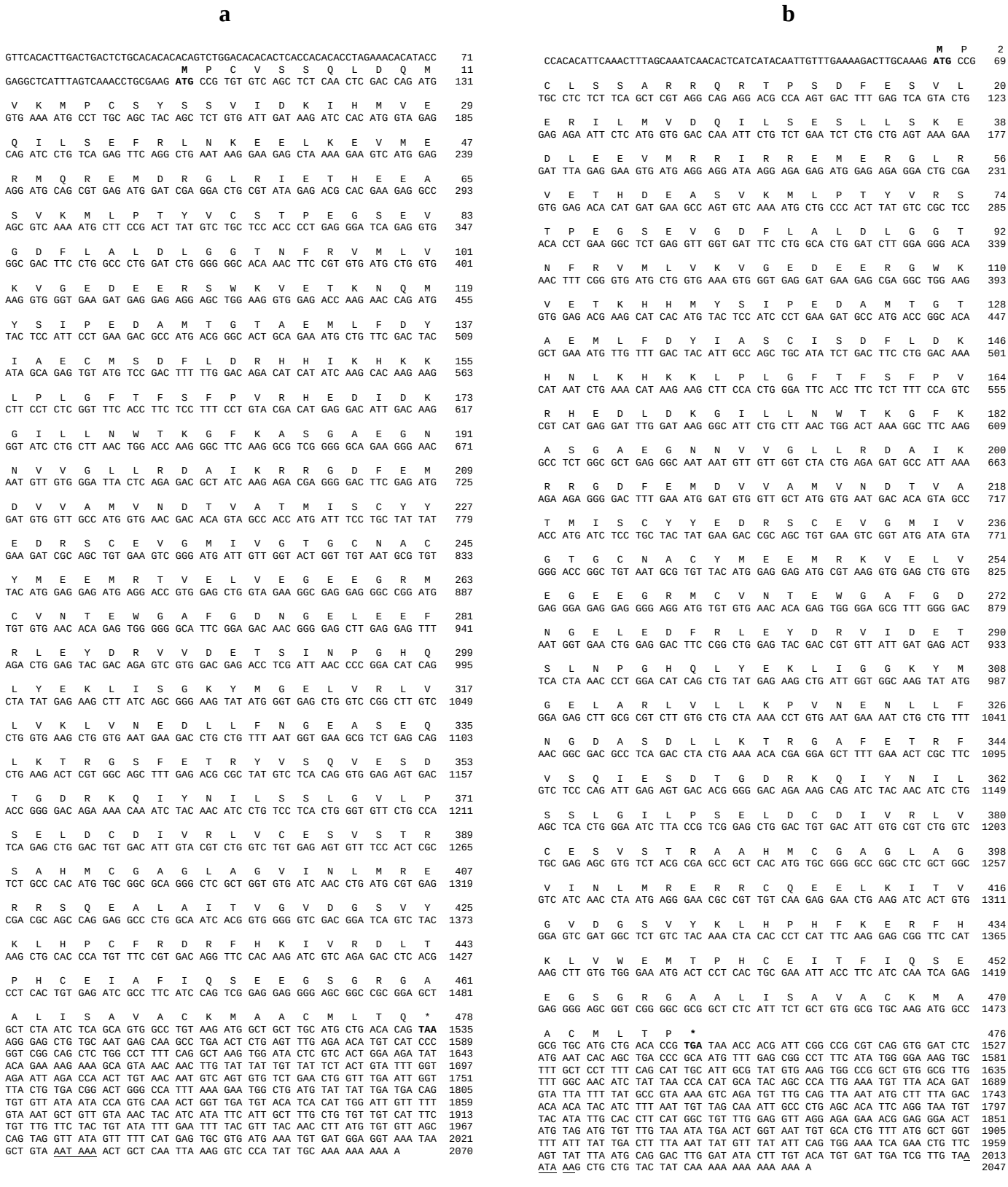


Figure 1

C

```
GTTTACAGTTCACTGACTCTGCATACACACACACTCATGCCAAACACACACTTACCACACACCTAGATACA    71
                                                  M   G   Q   M   G       5
CACGCGAGTGATTTAGAGAGCTAGACAAGATACTGTGCGTCAGCTCCTTC ATG GGC CAG ATG GGG    136

 K   M   P   C   S   L   S   S   V   L   E   R   V   I   M   V   E   Q     23
AAG ATG CCT TGT AGC CTC AGC TCT GTG CTA GAG AGA GTC ATC ATG GTG GAG CAG    190

 I   L   S   E   F   R   L   K   K   E   Q   L   K   E   V   M   K   R     41
ATC CTG TCG GAG TTC AGG CTG AAG AAG GAG CAG TTG AAG GAG GTG ATG AAG AGG    244

 M   M   R   E   M   D   R   G   L   R   V   E   T   H   Q   E   A   S     59
ATG ATG AGG GAG ATG GAC CGG GGA CTG CGT GTA GAG ACG CAC CAG GAG GCC AGT    298

 V   K   M   L   P   T   Y   V   C   S   T   P   E   G   S   E   V   G     77
GTC AAA ATG CTG CCC ACC TAC GTC TGT TCT ACC CCT GAA GGA TCA GAG GTG GGT    352

 D   F   L   A   L   D   L   G   G   T   N   F   R   V   M   L   V   K     95
GAT TTC CTG GCC CTG GAC CTG GGG GGG ACT AAC TTC CGT GTG ATG TTG GTG AAG    406

 V   G   E   D   E   E   R   G   W   K   V   E   T   K   H   Q   M   Y    113
GTG GGG GAG GAT GAG GAG AGG GGA TGG AAG GTG GAG ACC AAA CAC CAG ATG TAC    460

 S   I   S   E   D   A   M   T   G   T   A   E   M   L   F   D   Y   I    131
TCC ATC TCT GAG GAC GCA ATG ACA GGC ACG GCT GAG ATG CTC TTT GAC TAC ATT    514

 A   E   C   I   S   D   F   L   N   R   Q   H   I   K   H   K   K   L    149
GCT GAG TGT ATA TCA GAC TTC CTG AAC AGA CAA CAC ATC AAG CAC AAG AAG CTT    568

 P   L   G   F   T   F   S   F   P   V   R   H   E   N   I   D   K   G    167
CCT CTG GGT TTC ACC TTC TCT TTT CCT GTA CGA CAC GAG AAC ATA GAC AAG GGC    622

 I   L   L   N   W   T   K   G   F   K   A   S   G   A   E   G   N   N    185
ATC CTA CTG AAC TGG ACC AAA GGG TTC AAG GCG TCT GGA GCA GAG GGT AAC AAC    676

 V   V   G   L   L   R   D   A   I   K   R   R   G   D   F   E   M   D    203
GTG GTG GGA CTA CTG AGA GAT GCC ATC AAG AGG AGA GGG GAC TTT GAG ATG GAC    730

 V   V   A   M   V   N   D   T   V   A   T   M   I   S   C   Y   Y   E    221
GTC GTT GCC ATG GTG AAC GAT ACA GTT GCC ACC ATG ATA TCC TGT TAC TAT GAG    784

 D   R   S   C   E   V   G   M   I   V   G   T   G   C   N   A   C   Y    239
GAC CGC AGC TGC GAA GTG GGA ATG ATT GTG GGT ACT GGG TGT AAC GCT TGC TAC    838

 M   E   E   M   R   T   V   E   L   V   E   G   E   E   G   R   M   C    257
ATG GAG GAG ATG CGG ACA GTG GAG CTG GTG GAG GGG GAA GAG GGG AGG ATG TGT    892

 V   N   T   E   W   G   A   F   G   A   N   G   E   L   E   E   F   R    275
GTG AAC ACA GAG TGG GGG GCC TTT GGA GCC AAC GGA GAG CTG GAG GAG TTT AGA    946

 L   E   Y   D   R   V   V   D   E   T   S   L   N   P   G   Q   Q   L    293
CTG GAG TAC GAC AGG GTG GTG GAC GAG ACA TCA CTC AAC CCT GGA CAA CAA CTC   1000

 Y   E   K   L   I   S   G   K   Y   M   G   E   L   V   R   L   V   L    311
TAT GAA AAG CTG ATC AGT GGG AAG TAC ATG GGA GAG CTG GTG CGG CTG GTA TTG   1054

 L   K   L   V   N   E   E   L   L   F   N   G   E   A   S   D   L   L    329
TTG AAG CTG GTG AAC GAG GAG CTG CTG TTT AAC GGA GAA GCC TCT GAC CTG CTG   1108

 K   T   R   G   S   F   E   T   R   Y   V   S   Q   I   E   G   D   S    347
AAG ACT CGC GGC AGC TTT GAG ACG CGC TAC GTC TCC CAG ATA GAG GGT GAC TCT   1162

 G   D   R   K   Q   I   Y   N   I   L   S   T   L   G   V   L   P   S    365
GGA GAC AGG AAG CAG ATC TAC AAC ATC CTG TCT ACG CTG GGC GTG TTG CCG TCG   1216

 E   L   D   C   D   I   V   R   L   A   C   E   S   V   S   T   R   A    383
GAG CTG GAC TGT GAC ATA GTG CGT CTA GCT TGT GAG AGC GTG TCC ACG CGG GCA   1270

 A   H   M   C   G   A   G   L   A   G   V   I   N   R   M   R   E   R    401
GCA CAC ATG TGT GGG GCG GGG TTA GCC GGC GTC ATC AAC CGT ATG AGA GAA CGC   1324

 R   S   L   A   V   L   K   I   T   V   G   I   D   G   S   V   Y   K    419
CGC AGC CTG GCG GTG TTG AAG ATC ACT GTG GGC ATC GAC GGC TCC GTC TAC AAA   1378

 L   H   P   C   F   Q   D   R   F   H   K   V   V   R   E   L   T   P    437
CTC CAC CCC TGT TTC CAG GAC AGG TTC CAC AAA GTT GTG CGG GAG CTG ACG CCT   1432

 H   C   D   I   T   F   I   Q   S   E   E   G   S   G   R   G   A   A    455
CAC TGT GAC ATC ACC TTC ATC CAA TCA GAG GAA GGG AGT GGC CGG GGG GCG GCA   1486

 L   I   S   A   V   A   C   K   M   A   A   C   M   L   T   P   *        471
CTT ATC TCG GCA GTA GCC TGT AAG ATG GCA GCG TGC ATG CTG ACA CCC TGA GAG   1540
```

```
TCG TAA CAG CCT CCC ACC TAT AGG GGG CGA TGC AGG TGA ACC AGG CTT AAC TCT  1594
GAC CTC ATA CAG TAA ATT ATA CCT CAT CTA GGG TGA TCC AGC TCT GGC CTC CTC  1648
TTG AAG CTT TCA CAG AGT AAT CTG TAA CTC CTC ATT GGT GAG AGA AGC AAT ACG  1702
AGT GTG GAT AAC TCA CTG TAC ATA CAA ATA ACC CCA CCA CAC AGC ACA CAC AAC  1756
ATG GGA TAC ACA GAG ACA GAT TCA ATC ATC ACT TTT CCC AGT GTA CAG TAT TAT  1810
ACA GAT GTG TTT TTG AAC TGA TGA GCC TCT GAA AAT TCC CTG CAG TCC CAC AAT  1864
GCA TAA GTG GTG AGT CCC AGA AAG TAA CGA ACC CTT GTT TAA ACA TGT TCT CAA  1918
GCT TCC ACT GAT TGC AGG TAT TTA AAG CTT GAT TAA TAA TGT AAT ATA ATG TTA  1972
AAG GGA AAA AGA TTG GCA TTG ATT TAT AGC ATC AGT ATG TTT TTT AGT TGG ATG  2026
CTG TCT GTT TAA TCC AGA GAT GTT ACC GTA TAT AAA TTT AAA ACT AAG GGA TTA  2080
TCC CTT TCA AAT GTG TTG CAG TGC AGT CTT TCA CTT ACT GTT CGT GGG AAA TTA  2134
AAT ATT TAT TTG ATA GCA TAT TGA TAA AAG TGA TGT ATC TGT GAT GAG CGT AAA  2188
GGT CTT AAT TCC TAT GAT GTA TAT CTA AGA ATG ATA ATG TGC TTT GTA ATG TGA  2242
CCA TTT TTT ATT GTT GTA AGT AAA GTT GCT GTA ACA TAC#AAA TAT TTT TGT TGT  2296
GTG CCT ATC GTC ATT TGA AAT TGA AGG GGC CAA ATA ACT GAA GTG GCT TTA GCG  2350
TAA CTG AAG TGG CTT TAG CGT ACA TGA CCA CAT CCT GTC ACA AAT AAC TGT CAG  2404
GTT TCT GTT TCT GTT CAG TAA AGT GAG CAA CTC ACA CAA AGT TTC CCA GGC AAA  2458
CAC CCA ATG CAA CTG AAT AGG ACT GTA AAC TAC AGT CAT ACT CAG TGG TGT AAA  2512
GTA CCT CAG TAA AAA TGC TTG AAA GTA CTA CTT AAG TCG TTT TTT AGG GTA TCG  2566
GTA CTT TAC TTT ACT TAT ATA TAT TTT TGA CAA CTT TTA CTT TTA CTT CAC TAC  2620
ATT CCT AAT AAA AAT GAT GTA CTT TTA CGA AAA AAA AAA AAA A                2663
```

Figure 1bis

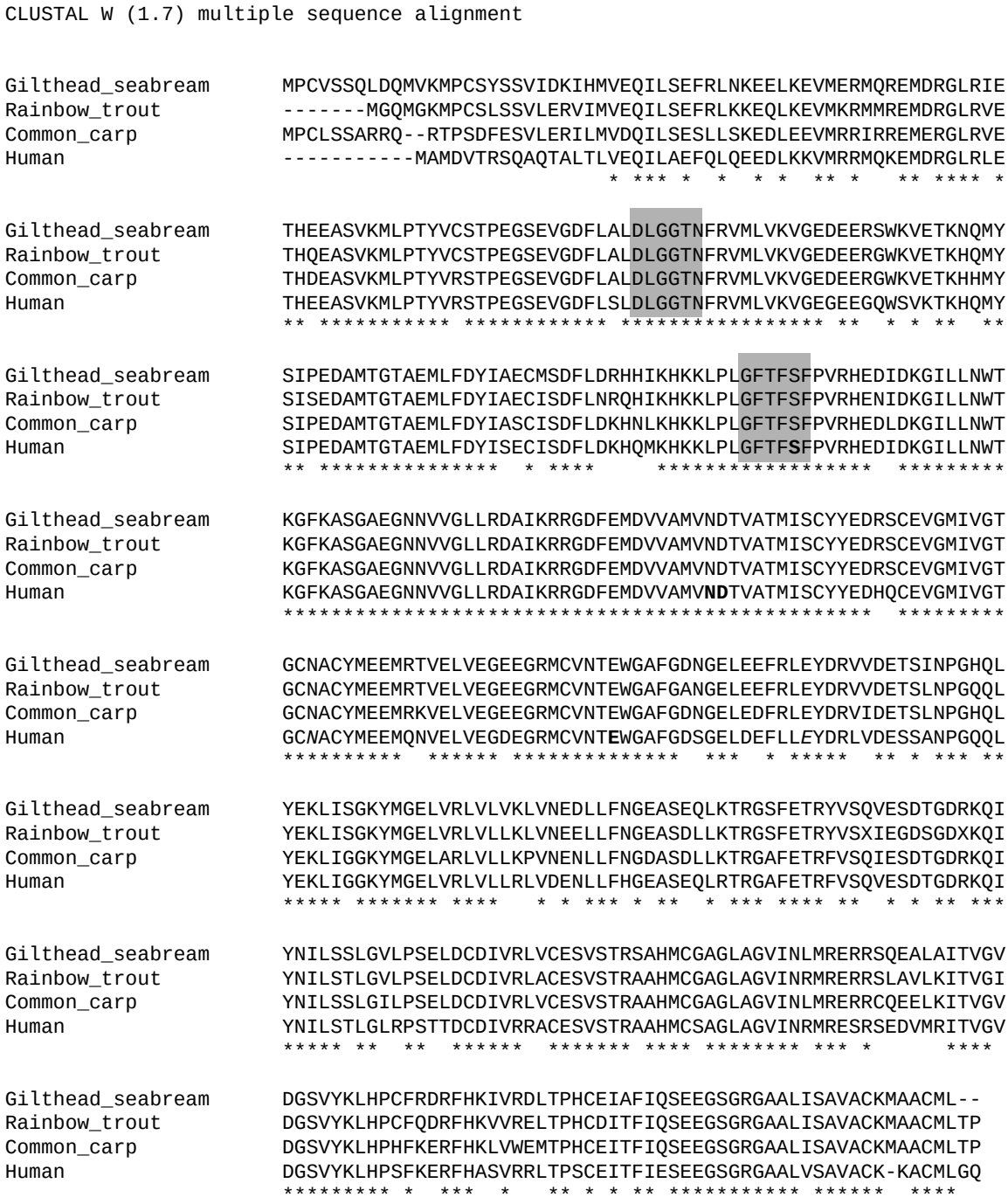
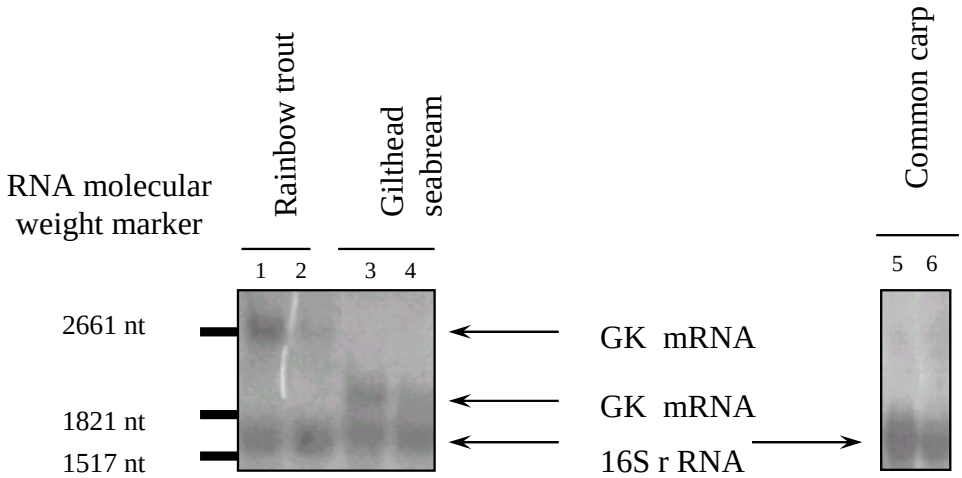

CLUSTAL W (1.7) multiple sequence alignment
Gilthead_seabream MPCVSSQLDQMVKMPCSYSSVIDKIHMVEQILSEFRLNKEELKEVMERMQREMDRGLRIE
Rainbow_trout -------MGQMGKMPCSLSSVLERVIMVEQILSEFRLKKEQLKEVMKRMMREMDRGLRVE
Common_carp MPCLSSARRQ--RTPSDFESVLERILMVDQILSESLLSKEDLEEVMRRIRREMERGLRVE
Human -----------MAMDVTRSQAQTALTLVEQILAEFQLQEEDLKKVMRRMQKEMDRGLRLE
Gilthead_seabream THEEASVKMLPTYVCSTPEGSEVGDFLALDLGGTNFRVMLVKVGEDEERSWKVETKNQMY
Rainbow_trout THQEASVKMLPTYVCSTPEGSEVGDFLALDLGGTNFRVMLVKVGEDEERGWKVETKHQMY
Common_carp THDEASVKMLPTYVRSTPEGSEVGDFLALDLGGTNFRVMLVKVGEDEERGWKVETKHHMY
Human THEEASVKMLPTYVRSTPEGSEVGDFLSLDLGGTNFRVMLVKVGEGEEGQWSVKTKHQMY
Gilthead_seabream SIPEDAMTGTAEMLFDYIAECMSDFLDRHHIKHKKLPLGFTFSFPVRHEDIDKGILLNWT
Rainbow_trout SISEDAMTGTAEMLFDYIAECISDFLNRQHIKHKKLPLGFTFSFPVRHENIDKGILLNWT
Common_carp SIPEDAMTGTAEMLFDYIASCISDFLDKHNLKHKKLPLGFTFSFPVRHEDLDKGILLNWT
Human SIPEDAMTGTAEMLFDYISECISDFLDKHQMKHKKLPLGFTFSFPVRHEDIDKGILLNWT
Gilthead_seabream KGFKASGAEGNNVVGLLRDAIKRRGDFEMDVVAMVNDTVATMISCYYEDRSCEVGMIVGT
Rainbow_trout KGFKASGAEGNNVVGLLRDAIKRRGDFEMDVVAMVNDTVATMISCYYEDRSCEVGMIVGT
Common_carp KGFKASGAEGNNVVGLLRDAIKRRGDFEMDVVAMVNDTVATMISCYYEDRSCEVGMIVGT
Human KGFKASGAEGNNVVGLLRDAIKRRGDFEMDVVAMVNDTVATMISCYYEDHQCEVGMIVGT
Gilthead_seabream GCNACYMEEMRTVELVEGEEGRMCVNTEWGAFGDNGELEEFRLEYDRVVDETSINPGHQL
Rainbow_trout GCNACYMEEMRTVELVEGEEGRMCVNTEWGAFGANGELEEFRLEYDRVVDETSLNPGQQL
Common_carp GCNACYMEEMRKVELVEGEEGRMCVNTEWGAFGDNGELEDFRLEYDRVIDETSLNPGHQL
Human GCNACYMEEMQNVELVEGDEGRMCVNTEWGAFGDSGELDEFLLEYDRLVDESSANPGQQL
Gilthead_seabream YEKLISGKYMGELVRLVLVKLVNEDLLFNGEASEQLKTRGSFETRYVSQVESDTGDRKQI
Rainbow_trout YEKLISGKYMGELVRLVLLKLVNEELLFNGEASDLLKTRGSFETRYVSXIEGDSGDXKQI
Common_carp YEKLIGGKYMGELARLVLLKPVNENLLFNGDASDLLKTRGAFETRFVSQIESDTGDRKQI
Human YEKLIGGKYMGELVRLVLLRLVDENLLFHGEASEQLRTRGAFETRFVSQVESDTGDRKQI
Gilthead_seabream YNILSSLGVLPSELDCDIVRLVCESVSTRSAHMCGAGLAGVINLMRERRSQEALAITVGV
Rainbow_trout YNILSTLGVLPSELDCDIVRLACESVSTRAAHMCGAGLAGVINRMRERRSLAVLKITVGI
Common_carp YNILSSLGILPSELDCDIVRLVCESVSTRAAHMCGAGLAGVINLMRERRCQEELKITVGV
Human YNILSTLGLRPSTTDCDIVRRACESVSTRAAHMCSAGLAGVINRMRESRSEDVMRITVGV
Gilthead_seabream DGSVYKLHPCFRDRFHKIVRDLTPHCEIAFIQSEEGSGRGAALISAVACKMAACML--
Rainbow_trout DGSVYKLHPCFQDRFHKVVRELTPHCDITFIQSEEGSGRGAALISAVACKMAACMLTP
Common_carp DGSVYKLHPHFKERFHKLVWEMTPHCEITFIQSEEGSGRGAALISAVACKMAACMLTP
Human DGSVYKLHPSFKERFHASVRRLTPSCEITFIESEEGSGRGAALVSAVACK-KACMLGQ

Figure 2

a


RNA molecular weight marker
Rainbow trout
Gilthead seabream
Common carp
1 2
3 4
5 6
2661 nt
1821 nt
1517 nt
GK mRNA
GK mRNA
16S r RNA

b

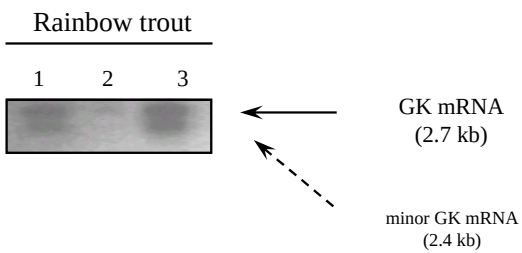

Rainbow trout
1 2 3
GK mRNA
(2.7 kb)
minor GK mRNA
(2.4 kb)

Figure 3

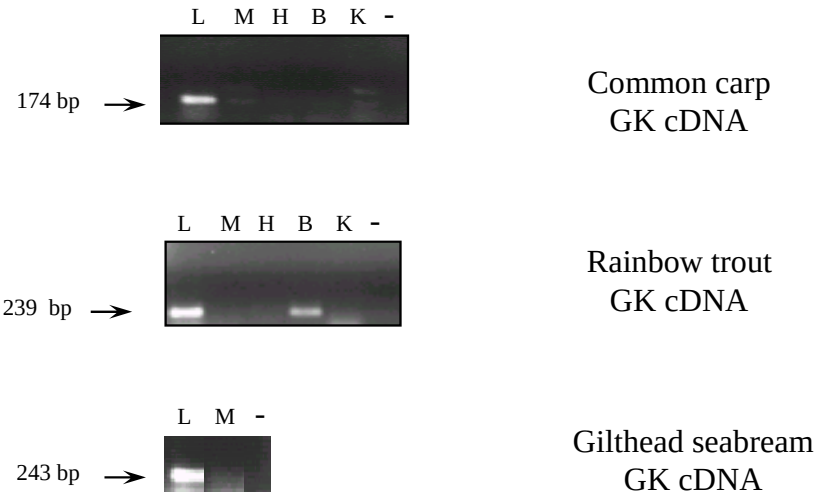

L M H B K -
174 bp
Common carp
GK cDNA
L M H B K -
239 bp
Rainbow trout
GK cDNA
L M -
243 bp
Gilthead seabream
GK cDNA

Figure 4

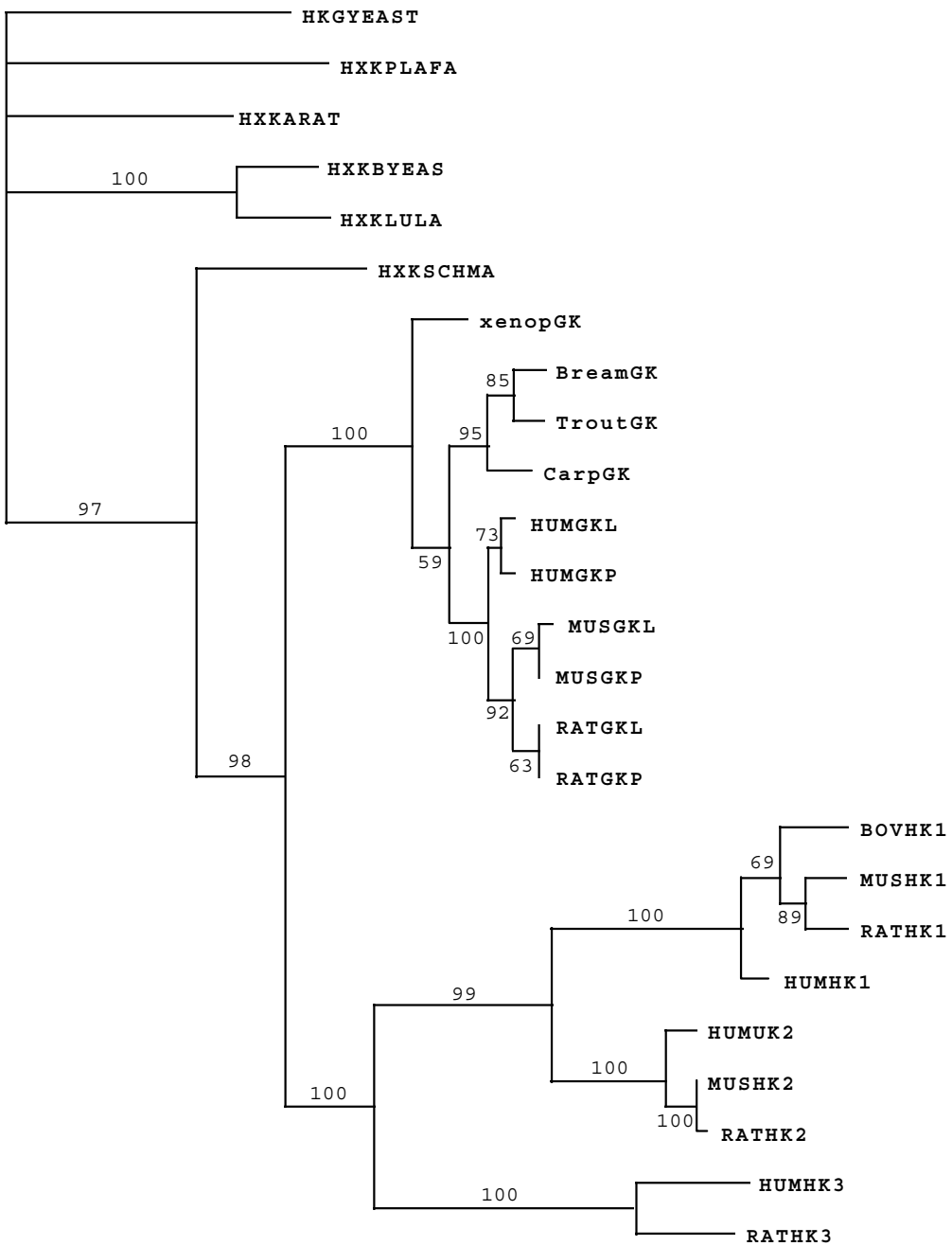

HKGYEAST
HXKPLAFA
HXKARAT
100
HXKBYEAS
HXKLULA
HXKSCHMA
xenopGK
85
BreamGK
TroutGK
100
95
CarpGK
97
HUMGKL
73
59
HUMGKP
100
69
MUSGKL
MUSGKP
92
RATGKL
98
63
RATGKP
BOVHK1
69
MUSHK1
100
89
RATHK1
HUMHK1
99
HUMUK2
100
MUSHK2
100
100
RATHK2
HUMHK3
100
RATHK3

Figure 5

Table 1. Primers used for the GK cDNA cloning by Race PCR and the RT-PCR analysis.

	Fish specific GK primers*	
	Race-PCR	RT-PCR
Gitlhead seabream	5' Race : (5'- TTCAGTAGGATGCCCTTGTC -3') 5' Race : (5'- GCAGTGCCCGTCATGGCGTC -3') 3' Race : (5'- TGTGATGCTGGTGAAGGTGG -3')	Forward : (5'- TGTGATGCTGGTGAAGGTGG -3') Reverse: (5'- TCATGTTGGTGAAGGTGGGG -3')
Rainbow trout	5' Race : (5'- TTCAGTAGGATGCCCTTGTC- -3') 5' Race : (5'- GCCGTGCCTGTCATTGCGTC -3') 3' Race : (5'- TCATGTTGGTGAAGGTGGGG -3')	Forward: (5'- TCATGTTGGTGAAGGTGGGG -3') Reverse: (5'- TTCAGTAGGATGCCCCTTGTC -3')
Common carp	5' Race : (5'- GGTTCCTATGTTTCAGATTA -3') 5' Race : (5'- GCTGTGCCGGTCATGGCATC -3') 3' Race : (5'- ACTGATGCTGGTCAAAGTGG -3')	Forward: (5'- ACTGATGCTGGTCAAAGTGG -3') Reverse: (5'- GGTTCTTATGTTTCAGATTA -3')

* : primers chose in the known fish GK clones [19]

Table 2. Origin of the HK amino acid sequences used in the phylogenetic analysis

name	size in amino acid	accession number	definition
HKGYEAST	500 AA	P17709	yeast GK (EC 2.7.1.2)
HXKPLAFA	493 AA	Q02155	Plasmodium HK (EC 2.7.1.1)
HXKARAT	435 AA	Q42525	arabidopsis thaliana HK (EC 2.7.1.1)
HXKBYEAST	486 AA	P04807	yeast HK B (PII) (EC 2.7.1.1)
HXKLULA	485 AA	P33284	kluyveromyces lactis HK (EC 2.7.1.1)
HXKSCHMA	451 AA	Q26609	Schistosoma mansoni HK (EC 2.7.1.1)
XENOPGK	458 AA	Q91754	Xenopus laevis GK
HUMGKL	466 AA	Q05810	Human HK type IV, liver isozyme (EC 2.7.1.1)
HUMGKP	465 AA	P35557	human HK type IV, pancreatic isozyme (EC 2.7.1.1)
MUSGKL	465 AA	P52791	mouse HK type IV, hepatic isozyme (EC 2.7.1.1)
MUSGKP	465 AA	P52792	mouse HK type IV, pancreatic isozyme (EC 2.7.1.1)
RATGKL	465 AA	P17711	rat HK type IV, hepatic isozyme (EC 2.7.1.1)
RATGKP	465 AA	P17712	rat HK type IV, pancreatic isozyme (EC 2.7.1.1)
BOVHK1	918 AA	P27595	bos taurus HK type I (EC 2.7.1.1)
MUSHK1	918 AA	P17710	mouse HK, type I (EC 2.7.1.1)
RATHK1	919 AA	P05708	rat HK, type I (EC 2.7.1.1)
HUMHK1	917 AA	P19367	human HK, type I (EC 2.7.1.1)
HUMUK2	917 AA	P52789	human HK, type II (EC 2.7.1.1)
MUSHK2	917 AA	O08528	rat HK, type II (EC 2.7.1.1)
RATHK2	917 AA	P27881	mouse HK, type II (EC 2.7.1.1)
HUMHK3	923 AA	P52790	human HK type III (EC 2.7.1.1)
RATHK3	924 AA	P27926	rat HK type III (EC 2.7.1.1)

Table 3. comparison between teleost and other vertebrate GK amino acid sequences by using Align program (Myers and Millers 1989)

Numbers are the percentage of identical amino acid residues between sequences. The highest value for each fish GK sequence is shown in bold. Sequences are from human (accession M90299), mouse (accession L38990), Xenopus laevis (accession X93494/1262840)

	Gitlhead seabream	Rainbow trout	Common carp	*Xenopus laevis*	Mouse	Human*
Gitlhead seabream	——	**88.3**	**85.8**	75.5	78.7	79.7
Rainbow trout		——	**85.1**	76.0	79.0	79.6
Common carp			——	76.3	77.5	78.8
Xenopus laevis				——	78.3	79.4
Mouse					——	93.8
Human						——

*: hepatic form