Two divergent leptin paralogues in zebrafish (*Danio rerio*) that originate early in teleostean evolution

Marnix Gorissen[1], Nicholas J. Bernier[2], Sander B. Nabuurs[3], Gert Flik[1] and Mark O. Huising[1,4]

[1]Department of Animal Physiology, Faculty of Science, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands.

[2]Department of Integrative Biology, University of Guelph, Guelph, Canada.

[3]Center for Molecular and Biomolecular Informatics (CMBI), Nijmegen Center for Molecular Life Sciences (NCMLS), Radboud University Nijmegen, Post Office Box 9101 6500 HB Nijmegen, Netherlands

[4]The Clayton Foundation Laboratories for Peptide Biology, The Salk Institute for Biological Studies, Peptide Biology Laboratories, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA.

Correspondence should be addressed to:

Marnix Gorissen

M.Gorissen@science.ru.nl

Short title: divergent leptin paralogues in zebrafish

Keywords: leptin, teleost, evolution, gene duplication

ABSTRACT:

We describe duplicate leptin genes in zebrafish (*Danio rerio*) that share merely 24% amino acid identity with each other and only 18% with human leptin. We were also able to retrieve a second leptin gene in medaka (*Oryzias latipes*). The presence of duplicate leptin genes in these two distantly related teleosts suggests that duplicate leptin genes are a common feature of teleostean fishes. Despite low primary sequence conservation, we are confident in assigning orthology between mammalian and zebrafish leptins for several reasons. Firstly, both zebrafish leptins share their characteristic gene structure and display key features of conserved synteny with mammalian leptin genes. Secondly, the cysteine residues that make up leptin's single disulphide bridge are equally spaced in mammals and zebrafish leptins and are unique among all members of the class-I helical cytokine family. Thirdly, the zebrafish leptins cluster with other fish leptins and mammalian leptins in phylogenetic analysis, supported by high bootstrap values. Within the leptin cluster, leptin-b forms a separate clade with the leptin-b orthologue from medaka. Finally, our prediction of the tertiary structures shows that both leptins conform to the typical four α-helix bundle structure of the class-I α-helical cytokines. The zebrafish leptins are differentially expressed; the liver shows high leptin-a expression (in concordance with what we observed for carp leptins), while leptin-b is expressed at much lower levels, which are downregulated further upon fasting. The finding of duplicate leptin genes in teleosts adds to our understanding of the evolution of leptin physiology in the early vertebrate lineage.

INTRODUCTION

The positional cloning of the *obese* (*ob*) gene in 1994 (Zhang, et al. 1994), identified the factor responsible for the morbid obesity of *ob/ob* mutant mice. This gene encodes a unique member of the class-I helical cytokine family, a 16 kDa protein named leptin after the Greek root *leptos* for lean. It is made up of a characteristic four α-helix bundle conformation (Zhang, et al. 1997). The key role of leptin in the regulation of body weight and energy homeostasis is well established (Morton, et al. 2006; Schwartz, et al. 2000). Leptin circulates in the bloodstream in proportion to the amount of body fat and signals to the brain. A major site of action is the *arcuate nucleus* (ARC), which contains two distinct populations of leptin-responsive neurons. One set co-expresses neuropeptide Y (NPY) and agouti-related protein (AgRP), is orexigenic and is inhibited by leptin (Broberger, et al. 1998), while the other expresses pro-opiomelanocortin (POMC) and cocaine and amphetamine regulated transcript (CART), is anorexigenic and is stimulated by leptin (Elias, et al. 1998).

Zhang and co-workers addressed the evolution of leptin by hybridizing genomic DNA of vertebrates that originated early in vertebrate evolution, including teleost fish, with a murine *ob* probe (Zhang et al. 1994). Positive signals from teleost genomic DNA led them to conclude that leptin is highly conserved throughout the vertebrates. Despite the detection of leptin-like immunoreactivity in the blood and liver it took more than a decade to characterize the first teleost leptin orthologue (Huising, et al. 2006a; Kurokawa, et al. 2005) or even amphibian leptin orthologues (Boswell, et al. 2006; Crespi and Denver 2006). No bona-fide avian and reptilian leptin genes have been described to date (Huising, et al. 2006b). Both fish and *Xenopus* leptin show a low degree of primary sequence conservation compared to human (varying from 13 to 30% amino acid identity, respectively). Although the mere presence of a leptin orthologue in teleost fish supports the notion of leptin's evolutionary conservation, leptin is among the class-I helical cytokines with the poorest sequence conservation throughout the vertebrate subphylum

(Huising et al. 2006b). In fish, a major site of leptin expression is the liver (Huising et al. 2006a; Kurokawa et al. 2005), which is rich in fat droplets and has therefore been suggested an appropriate site to monitor adipose stores. Yet, our understanding of the contribution of leptin to the regulation of energy metabolism in fish is scant and a key role of leptin in the regulation of body weight and energy homeostasis in non-mammalian vertebrates has not been established thus far (Gorissen, et al. 2006; Volkoff, et al. 2005). In carp, hepatic leptin mRNA increases postprandially, but not after fasting or feeding to satiation for up to six weeks (Huising et al. 2006a).

It is well known that teleost fish possess duplicate copies for a number of genes (Taylor, et al. 2003; Volff 2005). Therefore, we searched the zebrafish genome database to see if leptin too occurs in duplicate. Here, we demonstrate duplicate leptin genes in zebrafish (*Danio rerio*). An earlier systematic search of the zebrafish genome database revealed a predicted leptin gene with high (61-62%) amino acid identity to both carp leptin-a I and leptin-a II (accession number BN000830) now designated leptin-a (Huising et al. 2006a). We cloned this leptin gene and a second, substantially different and paralogous leptin gene in zebrafish. Both zebrafish leptin paralogues share 24% primary amino acid sequence identity with each other and 18% with mammalian leptins. Zebrafish leptin-a shares high primary sequence conservation with both carp leptins (61-62%); leptin-b, however, shares only 25% amino acid identity with both carp leptins. Despite these low identities, conservation of gene structure, tertiary structure, stable phylogenetic analysis and synteny substantiate the unambiguous orthology of zebrafish leptin-a and leptin-b with mammalian leptins.

MATERIALS AND METHODS

Animals

Zebrafish (*Danio rerio*) were commercially obtained and reared in two litre tanks at 26°C with recirculating, UV-treated, Nijmegen tap water. Eight fish were kept in a single aquarium and fed 2.5% body weight Tetra-min (Tetra, Melle Germany) each day. Eight other fish were not fed for two weeks. For the determination of leptin tissue distribution, fish were fed 2.5% body weight daily and sacrificed one hour after feeding. All fish were euthanized in a 0.1% (w/v) 2-Phenoxyethanol solution. Animal experiments were performed in accordance with national legislation and approved by the ethical committee of the Radboud University Nijmegen.

Identification of zebrafish leptin paralogues

We screened the ENSEMBL zebrafish genome (www.ensembl.org) with several teleost leptin sequences, using the BLAST algorithm (Altschul, et al. 1997). The initial screen revealed two leptin-like sequences, one of which was already predicted in an earlier screen of the zebrafish genome (third party annotation (TPA) accession number: BN000830; (Huising et al. 2006a)). Using primers zf.leptin-a.fw, leptin-a.rv and leptin-b.fw, leptin-b.rv (table 1), based on these partial leptin sequences, two cDNA sequences were obtained from the liver and gonads, respectively. RNA isolation, cDNA synthesis, cloning and sequencing was carried out as previously described (Metz, et al. 2005). Briefly, PCR products were ligated and cloned in TOP10 chemically competent *E. coli* in the pCR4-TOPO vector (Invitrogen, Carlsbad, CA, USA). Plasmid DNA was isolated with a miniprep kit (BioRad, Hercules, USA) and sequences were determined from both strands using the ABI prism big dye terminator cycle sequencing ready reaction kit (Applied Biosystems, Foster City, USA).

109

110 *Phylogenetic analysis*

111 Multiple sequence alignments were carried out using ClustalW
112 (http://www.ebi.ac.uk/Tools/clustalw/; (Thompson, et al. 1994)). A phylogenetic tree was
113 constructed based on amino acid difference (p-distance) with the neighbour-joining algorithm
114 (pairwise deletion) in MEGA version 3.1 (Kumar, et al. 2004). The reliability of the tree was
115 assessed by bootstrapping, using 1000 replications. Only full-length coding sequences were used
116 for analysis.

117 In order to determine synteny between the zebrafish leptin paralogues and human leptin, we
118 mapped the upstream and downstream genes of leptin on the respective chromosomes of
119 zebrafish and human using the ENSEMBL genome browser (www.ensembl.org).

120

121 *Modelling of tertiary structures*

122 The structure of human leptin (PDB entry 1AX8), which was resolved at 2.4 Å resolution (Zhang
123 et al. 1997), was used as a template to build models of zebrafish leptin-a and leptin-b. Initial
124 alignments were obtained from the PSIPRED fold recognition server (McGuffin and Jones
125 2003). Side-chain rotamers were modeled using SCWRL3.0 (Canutescu, et al. 2003). Both models
126 were refined in YASARA using the YAMBER2 forcefield (Krieger, et al. 2004). Coordinate files
127 are available from the authors on request.

128

129 *Expression of zebrafish leptins*

130 Relative expression of zebrafish leptin paralogues was assessed by real-time qPCR. We designed
131 primers using primer express software (Table 1; Applied Biosystems). Five µl cDNA and 300 nM

132 forward and reverse primers were added to 12.5 µl SYBR Green mastermix (Applied biosystems).
133 The total volume was adjusted to 25 µl with deionised H_2O. qPCR (ten minutes 95°C, 40 cycles
134 of 15 seconds 95°C and one minute 60°C) was carried out using a GeneAmp 7500 sequence
135 detection system (Applied Biosystems). Different samples were run on a single plate. Dual
136 internal standards (40S ribosomal protein S11 and β-actin) were incorporated in all measurements
137 and results were confirmed to be very similar following standardisation to either gene. Only
138 results relative to 40S are shown. Constitutive expression of leptin in zebrafish organs and tissues
139 was corrected for primer efficiency and plotted as a ratio between target gene *vs.* reference gene.
140 Relative expression of leptin paralogues in the liver following fasting was corrected for primer
141 efficiency and reference gene, and plotted relative to controls.

142

143 RESULTS

144 *Zebrafish expresses duplicate and divergent leptin genes*

145 A systematic BLAST search of the Ensembl zebrafish genome database with mammalian leptin
146 sequences revealed two partial leptin sequences, one of which represented leptin-a (already
147 described by Huising et al, 2006a), the other represented a new leptin-like orthologue that we
148 named leptin-b. The (automated) genomic sequences were corrected by hand for correct splice
149 sites and the obtained sequences were used in a homology cloning approach to identify both
150 leptin cDNA sequences. Protein-protein BLAST (BLASTp) showed significant hits with other
151 fish leptins (**table 2**). The cDNA- and deduced amino acid sequences of zebrafish leptins are
152 shown in **figure 1**. Both leptin-a and leptin-b are comparable in size, 166 and 168 amino acids
153 respectively, both with a predicted signal peptide of 20 amino acids. Previously, we described two
154 highly similar leptin genes in common carp (Huising et al. 2006a) which we designated leptin-I
155 and leptin-II. These carp leptin paralogues are likely the result of the recent genome duplication

~16 Mya that led to the tetraploidization of the common carp genome (Larhammar and Risinger 1994). Our results suggest that zebrafish leptin-a and leptin-b are the result of the ancient genome duplication that teleost fish experienced (Taylor et al. 2003; Volff 2005). Therefore, we amend the names of the previously described carp leptins to leptin-a-I and leptin-a-II. We want to stress that the low amino acid identity of the leptin proteins between fish and mammals serves as a reminder that we assign the name leptin solely based on the structural similarities described above. Orthologous proteins do not by default share analogous roles, particularly proteins that share so little of their primary amino acid sequences as teleostean and mammalian leptins do (Huising et al. 2006a).

Characteristics of zebrafish leptins

The amino acid identity between zebrafish leptin-a and leptin-b is 24%. Zebrafish leptin-a is more similar (60% primary amino acid sequence identity) to carp leptin-a-I and leptin-a-II. The identity between zebrafish leptin-b and carp leptins is at 25% only marginally higher than the identity between leptin-b and mammalian leptins (19%; **figure 2**, **table 3**). The cysteine residues that make up leptin's single disulphide bridge, connecting the carboxy-terminal ends of α-helices C and D are conserved. Both zebrafish leptin genes are encoded by two exons that are similar in size compared to mammalian leptins (**figure 3**). Zebrafish leptin genes possess a short intron, with consensus 5' donor (gt) and 3' acceptor (ag) splice sites. The intron phase indicates whether the intron is situated in between triplets (phase 0), or following the first or second base of a triplet (phase 1 or phase 2 respectively. The intron phase for both zebrafish leptins is identical to the intron phase of mammalian leptins: phase 0.

Our models of both zebrafish leptins conform to the typical four α-helix conformation (up-up-down-down) of human leptin (**figure 4**), indicating that the tertiary structures of zebrafish leptins

are comparable to mammalian leptins. In contrast to all other leptin sequences, leptin-b contains an additional cysteine residue in helix D. From the position of this cysteine (indicated in red in figure 4) we cannot draw firm conclusions regarding the availability of this cysteine to form intermolecular disulphide bridges.

Phylogeny of zebrafish leptins

The zebrafish leptin paralogues cluster together with other vertebrate leptin genes, supported by a high bootstrap value (98), supporting the orthology of both zebrafish leptins with mammalian leptins (**figure 5**). Within the leptin cluster, the overall topology of the phylogenetic tree adheres to the established pattern of evolution, as the teleost leptin cluster branches off before the separation of the amphibian and mammalian cluster. Within the mammalian leptin cluster, the only known sequence of a marsupial leptin (that of the fat-tailed dunnart) branches outside the leptin sequences of placental mammals. In the teleost leptin cluster, zebrafish leptin-a and the carp leptins form a separate clade. We also screened other fish databases in order to assess the presence of leptin-b orthologues in other teleost fish species. Using the zebrafish leptin-b sequence in a BLAST search of the ENSEMBL medaka (*Oryzias latipes*) genome, we retrieved a leptin-b orthologue with 28% amino acid identity to zebrafish leptin-b (BN001183). Zebrafish and medaka leptin-b form a separate clade within the teleost leptin cluster.

Both zebrafish leptin genes share synteny with human leptin

To further substantiate the orthology of the zebrafish leptin paralogues to mammalian leptins, we compared the synteny of both zebrafish leptins with human leptin. Synteny refers to the order and orientation of the genes of a chromosomes and tends to be a conserved feature across

203 species. For each zebrafish leptin, several genes are found in synteny with mammalian leptin
204 (**figure 6**). The leptin-a gene of zebrafish is located next to RNA binding motif 28 (RBM28), as is
205 the human leptin gene. In close proximity of leptin-b, Staphylococcal nuclease domain-containing
206 protein 1 (SND1) and GRIP and coiled-coil domain containing 1 (GCC1) are found – again,
207 these are also found in close proximity of human leptin.

208

209 *Constitutive expression of zebrafish leptins*

210 Zebrafish leptins (**figure 7**) show a differential expression pattern. Whereas leptin-a is
211 prominently expressed in the liver, in accordance with previous observations of carp leptin-a-I
212 and -II, leptin-b is not. Leptin-a is expressed at higher levels than leptin-b in most organs except
213 the ovary, which is a major site of leptin-b mRNA expression.

214

215 *Leptin mRNA expression after fasting for one week*

216 To gain insight in possible physiological functions of the leptin paralogues, we investigated leptin
217 mRNA expression after fasting for one week. Leptin-a mRNA levels show no significant
218 response to fasting for one week (**figure 8**). In contrast, hepatic leptin-b expression is
219 significantly downregulated ($P<0.05$) after one week of food deprivation.
220

221 DISCUSSION

222 Zebrafish possesses duplicate leptin genes, coding for leptin-a and leptin-b, that differ
223 substantially from each other (24% amino acid identity). It is possible that a major genome
224 duplication that took place ~300 Mya in the early fish lineage (Taylor et al. 2003; Volff 2005)

225 resulted in duplicated leptins. The discovery of a leptin-b orthologue in the Japanese medaka
226 supports this view as zebrafish and medaka represent two distant teleost lineages, the *Cypriniformes*
227 and the *Beloniformes* respectively, that shared their last common ancestor ~296 Mya (Hoegg and
228 Meyer 2005); **figure 9**). In contrast, from the primary sequence identity and phylogenetic analysis,
229 it follows that the duplicate carp leptins that we described recently (Huising et al. 2006a) likely
230 resulted from the more recent genome duplication in carp (~16 Mya; Larhammer and Risinger,
231 1994) and represent the duplicated orthologues of zebrafish leptin-a. Therefore, we propose that
232 these carp leptin sequences should be renamed leptin-a-I and leptin-a-II. This observation,
233 combined with the identification of leptin-b in two distantly related fish substantiates the view
234 that more bony fishes express orthologues of leptin-b. Gene duplications, and genome
235 duplications in particular, are considered the main thrust contributing to the expansion of an
236 organism's gene repertoire, as the presence of newly duplicated paralogues allows one of the two
237 paralogues of a pair to drift and on occasion acquire a novel function while the original function
238 is maintained by the other. Gene duplications in the teleost lineage are common, and there are
239 several well-documented examples of large scale (often referred to as whole) genome duplication
240 events. A major genome duplication (Taylor et al. 2003; Volff 2005) is thought to have yielded
241 several duplicate class-I helical cytokines, *viz.* duplicate interleukin-11(Huising, et al. 2005), IL-
242 12p35 (Huising et al. 2006b), CXCL12 (Huising, et al. 2004) and cytokine receptor (IL12p40;
243 (Huising, et al. 2006a) genes. We could not retrieve a leptin-b orthologue from the available
244 pufferfish genomes (tiger pufferfish; *Takifugu rubripes*, and the green spotted pufferfish; *Tetraodon*
245 *nigroviridis*). While one reason for our inability to retrieve leptin-b orthologues from these species
246 may be that their respective genomes are incomplete, it is also possible that the *Tetraodontoformes*
247 may have lost leptin-b from their gene repertoire. In the genome of *T. nigroviridis*, we found two
248 regions with a conserved genomic neighbourhood compared to human leptin. Indeed, only one

249 of these loci carries a leptin orthologue, which is strong support for the hypothesis that the
250 pufferfish lineage does not possess duplicate leptin genes.

251 Recently, multiple entries have been submitted in the EMBL database for several fish leptin
252 orthologues that all share 97-99% sequence similarity at the nucleotide level (AY497007,
253 AY547279, AY547322, AY551335, AY551336, AY551337, AY551338, AY551339, AY551340,
254 DQ784814, DQ784815, DQ784816). Non-synonymous substitutions are subject to selection as
255 they result in differences in amino acid sequence, whereas synonymous substitutions are generally
256 not. Therefore, the almost complete absence of synonymous substitutions (over 97% nucleotide
257 identity) between these deposited 'teleost' leptin sequences and mammalian leptin sequences
258 would represent an extraordinary and very unlikely example of evolutionary convergence, as
259 teleosts and mammals shared their last common ancestor over 450 million years ago. Instead
260 these sequences should be regarded as artefacts. A similar situation unfortunately has occurred
261 for chicken leptin, that was reported to be highly similar to mouse leptin by two independent
262 groups (Ashwell, et al. 1999; Taouis, et al. 1998). Subsequent studies have raised concerns
263 regarding the validity of these published chicken leptin sequences (Doyon, et al. 2001; Friedman-
264 Einat, et al. 1999; Huising et al. 2006b; Sharp, et al. 2008).

265 Despite the relatively low amino acid conservation that was previously noted for other teleost
266 leptins, we are confident to assign orthology between zebrafish leptin-b and mammalian leptins,
267 supported by several key features of zebrafish leptin-b. First, both zebrafish leptin genes are
268 encoded by two exons of comparable size to the ones coding for mammalian leptins. Vertebrate
269 class-I cytokines are typically encoded by three or more (usually five) exons. In fact, the only
270 class-I helical cytokine other than leptin composed of two exons is ciliary neurotrophic factor
271 (CNTF) (Huising et al. 2006b), which differs substantially in primary sequence as well as gene
272 structure from leptin. Furthermore, the spacing of the two cysteine residues that make up leptin's
273 single disulphide bridge is unique among class-I helical cytokines (Huising et al. 2006b). Thirdly,

274 the stable phylogenetic clustering of the zebrafish leptin sequences with other fish leptins, as with
275 the mammalian leptins supports the unambiguous identity of the two zebrafish leptins. Finally,
276 the predicted tertiary structure of zebrafish leptin-b, conforming to the human crystal structure
277 of leptin, and the conservation of synteny between the mammalian leptin-locus and both
278 zebrafish leptin loci further strengthens the assignment of orthology between zebrafish leptins
279 and mammalian leptins.

280 An intriguing feature of the leptin-b sequence is the cysteine residue at the N-terminus of α-helix
281 D. We designed 3D models of leptin-b to address the spatial orientation of this additional
282 cysteine residue to see if this free cysteine would potentially be surface-exposed – and thus
283 available for disulphide bridging – or is buried within the leptin's hydrophobic core. These
284 models did not allow a firm prediction of the availability of this cysteine to form disulphide
285 bridges, either within one leptin molecule or between two molecules because its position in the
286 models is at the boundary of the protein surface and the protein core. It is possible that the
287 residue is buried within the protein, and as a result not exposed to the environment and not
288 available for disulphide interactions. The predicted mature leptin-b peptide contains no cysteine
289 to form a disulphide bridge with the helix-D cysteine. A similar phenomenon has been observed
290 for interleukin-11 genes in teleosts. Fish IL-11a and IL-11b both possess a single cysteine residue
291 near the C-terminus, whereas mammalian IL-11 does not (Huising et al. 2005). Medaka leptin-b
292 lacks an additional cysteine, indicating that this is not a universal feature among teleostean leptin-
293 b genes. The elucidation of additional teleost leptin-b sequences will shed light on the uniqueness
294 of this characteristic of zebrafish leptin-b.

295 We observed substantial differences between the expression patterns of zebrafish leptin-a and
296 leptin-b. It is now generally accepted that leptin, in addition to its 'classical' role is truly
297 pleiotropic (De Rosa, et al. 2007; Popovic, et al. 2001). Indeed in zebrafish, leptin-a and leptin-b
298 are expressed in considerable amounts in the pituitary gland. We do not know the exact nature of

299 the pituitary cells that (co-) express leptin in fish, nor the exact function of this leptin; in
300 mammals it is known that leptin is expressed in around 70% of the corticotropes and to a lesser
301 extent in somatotropes (21%), gonadotropes (29-33%), and thyrotropes (32%) (Popovic et al.
302 2001). We propose that fish leptin produced in the pituitary gland must have additional, local
303 (paracrine?) functions that allow zebrafish to maintain equilibrium in the face of challenges to
304 homeostasis.

305 Whereas the high level of expression of leptin-a in the zebrafish liver conforms to the expression
306 pattern observed for carp leptins, leptin-b is expressed at lower levels in the liver. Interestingly, it
307 is this hepatic leptin-b mRNA level that decreases after fasting. The sheer size of the fish liver
308 may guarantee a sufficient output of leptin(-b) protein, despite the relatively low leptin-b mRNA
309 expression level.

310 Leptin-b shows highest expression in the ovaries, which hardly express leptin-a. In mammals,
311 leptin serves a function in the regulation of reproduction as *ob/ob* mice treated with leptin recover
312 fertility (Archanco, et al. 2003; Caprio, et al. 2001). Given the high expression of leptin-b in
313 zebrafish ovaries, the reproductive function of leptin in this species may be carried out by leptin-
314 b.

315 In addition to the marked differences in leptin's primary sequences between teleosts and
316 mammals – which indicates potential differences in function – we now have demonstrated the
317 existence of a second, equally divergent leptin in zebrafish and medaka that is likely a feature
318 shared by more teleost fishes. The future challenge will be to unravel the physiological function
319 of both leptin genes. In fact, the presence of two highly divergent orthologues of mammalian
320 leptin in bony fish is testimony to the dynamic evolutionary history of leptin as it suggests the
321 possibility of a redundant leptin network in teleosts. Furthermore, it adds fuel to the proposition

322 that fish leptins, acting redundantly or independently, have acquired fundamentally different roles
323 compared to mammalian leptins.

324

ACKNOWLEDGEMENTS

We thank Dr. Paul Kievit for advice regarding genome databases and Mr. F.A. Tom Spanings for excellent fish husbandry.

DISCLOSURE

The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work. The authors did not receive additional funds for the research described in this paper.

335 REFERENCES

336 Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ 1997 Gapped
337 BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*
338 **25** 3389-3402.
339 Archanco M, Muruzabal FJ, Llopiz D, Garayoa M, Gomez-Ambrosi J, Fruhbeck G & Burrell
340 MA 2003 Leptin expression in the rat ovary depends on estrous cycle. *J Histochem Cytochem* **51**
341 1269-1277.
342 Ashwell CM, Czerwinski SM, Brocht DM & McMurtry JP 1999 Hormonal regulation of leptin
343 expression in broiler chickens. *Am J Physiol* **276** R226-232.
344 Boswell T, Dunn IC, Wilson PW, Joseph N, Burt DW & Sharp PJ 2006 Identification of a non-
345 mammalian leptin-like gene: characterization and expression in the tiger salamander (Ambystoma
346 tigrinum). *Gen Comp Endocrinol* **146** 157-166.
347 Broberger C, Johansen J, Johansson C, Schalling M & Hokfelt T 1998 The neuropeptide
348 Y/agouti gene-related protein (AGRP) brain circuitry in normal, anorectic, and monosodium
349 glutamate-treated mice. *Proc Natl Acad Sci U S A* **95** 15043-15048.
350 Canutescu AA, Shelenkov AA & Dunbrack RL, Jr. 2003 A graph-theory algorithm for rapid
351 protein side-chain prediction. *Protein Sci* **12** 2001-2014.
352 Caprio M, Fabbrini E, Isidori AM, Aversa A & Fabbri A 2001 Leptin in reproduction. *Trends*
353 *Endocrinol Metab* **12** 65-72.
354 Crespi EJ & Denver RJ 2006 Leptin (ob gene) of the South African clawed frog Xenopus laevis.
355 *Proc Natl Acad Sci U S A* **103** 10092-10097.
356 De Rosa V, Procaccini C, Cali G, Pirozzi G, Fontana S, Zappacosta S, La Cava A & Matarese G
357 2007 A key role of leptin in the control of regulatory T cell proliferation. *Immunity* **26** 241-255.
358 Doyon C, Drouin G, Trudeau VL & Moon TW 2001 Molecular Evolution of Leptin. *Gen Comp*
359 *Endocrinol* **124** 188-198.
360 Elias CF, Lee C, Kelly J, Aschkenasi C, Ahima RS, Couceyro PR, Kuhar MJ, Saper CB &
361 Elmquist JK 1998 Leptin activates hypothalamic CART neurons projecting to the spinal cord.
362 *Neuron* **21** 1375-1385.
363 Friedman-Einat M, Boswell T, Horev G, Girishvarma G, Dunn IC, Talbot RT & Sharp PJ 1999
364 The Chicken Leptin Gene: Has It Been Cloned? *Gen Comp Endocrinol* **115** 354-363.
365 Gorissen M, Flik G & Huising MO 2006 Peptides and proteins regulating food intake: a
366 comparative view. *Animal Biology* **56** 447-473.
367 Hedges SB 2002 The origin and evolution of model organisms. *Nat Rev Genet* **3** 838-849.
368 Hoegg S & Meyer A 2005 Hox clusters as models for vertebrate genome evolution. *Trends Genet*
369 **21** 421-424.
370 Huising MO, Geven EJ, Kruiswijk CP, Nabuurs SB, Stolte EH, Spanings FA, Verburg-van
371 Kemenade BM & Flik G 2006a Increased leptin expression in common Carp (Cyprinus carpio)
372 after food intake but not after fasting or feeding to satiation. *Endocrinology* **147** 5786-5797.
373 Huising MO, Kruiswijk CP & Flik G 2006b Phylogeny and evolution of class-I helical cytokines.
374 *J Endocrinol* **189** 1-25.
375 Huising MO, Kruiswijk CP, van Schijndel JE, Savelkoul HF, Flik G & Verburg-van Kemenade
376 BM 2005 Multiple and highly divergent IL-11 genes in teleost fish. *Immunogenetics* **57** 432-443.
377 Huising MO, van der Meulen T, Flik G & Verburg-van Kemenade BM 2004 Three novel carp
378 CXC chemokines are expressed early in ontogeny and at nonimmune sites. *Eur J Biochem* **271**
379 4094-4106.
380 Huising MO, van Schijndel JE, Kruiswijk CP, Nabuurs SB, Savelkoul HF, Flik G & Verburg-van
381 Kemenade BM 2006a The presence of multiple and differentially regulated interleukin-12p40
382 genes in bony fishes signifies an expansion of the vertebrate heterodimeric cytokine family. *Mol*
383 *Immunol* **43** 1519-1533.

Krieger E, Darden T, Nabuurs SB, Finkelstein A & Vriend G 2004 Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins* **57** 678-683.
Kumar S, Tamura K & Nei M 2004 MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5** 150-163.
Kurokawa T, Uji S & Suzuki T 2005 Identification of cDNA coding for a homologue to mammalian leptin from pufferfish, Takifugu rubripes. *Peptides* **26** 745-750.
Larhammar D & Risinger C 1994 Molecular genetic aspects of tetraploidy in the common carp Cyprinus carpio. *Mol Phylogenet Evol* **3** 59-68.
McGuffin LJ & Jones DT 2003 Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics* **19** 874-881.
Metz JR, Geven EJ, van den Burg EH & Flik G 2005 ACTH, alpha-MSH, and control of cortisol release: cloning, sequencing, and functional expression of the melanocortin-2 and melanocortin-5 receptor in Cyprinus carpio. *Am J Physiol Regul Integr Comp Physiol* **289** R814-826.
Morton GJ, Cummings DE, Baskin DG, Barsh GS & Schwartz MW 2006 Central nervous system control of food intake and body weight. *Nature* **443** 289-295.
Popovic V, Damjanovic S, Dieguez C & Casanueva FF 2001 Leptin and the pituitary. *Pituitary* **4** 7-14.
Schwartz MW, Woods SC, Porte D, Jr., Seeley RJ & Baskin DG 2000 Central nervous system control of food intake. *Nature* **404** 661-671.
Sharp PJ, Dunn IC, Waddington D & Boswell T 2008 Chicken leptin. *Gen Comp Endocrinol* **158** 2-4.
Taouis M, Chen JW, Daviaud C, Dupont J, Derouet M & Simon J 1998 Cloning the chicken leptin gene. *Gene* **208** 239-242.
Taylor JS, Braasch I, Frickey T, Meyer A & Van de Peer Y 2003 Genome duplication, a trait shared by 22000 species of ray-finned fish. *Genome Res* **13** 382-390.
Thompson JD, Higgins DG & Gibson TJ 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22** 4673-4680.
Volff JN 2005 Genome evolution and biodiversity in teleost fish. *Heredity* **94** 280-294.
Volkoff H, Canosa LF, Unniappan S, Cerda-Reverter JM, Bernier NJ, Kelly SP & Peter RE 2005 Neuropeptides and the control of food intake in fish. *Gen Comp Endocrinol* **142** 3-19.
Zardoya R & Doadrio I 1999 Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. *J Mol Evol* **49** 227-237.
Zhang F, Basinski MB, Beals JM, Briggs SL, Churgay LM, Clawson DK, DiMarchi RD, Furman TC, Hale JE, Hsiung HM, et al. 1997 Crystal structure of the obese protein leptin-E100. *Nature* **387** 206-209.
Zhang Y, Proenca R, Maffei M, Barone M, Leopold L & Friedman JM 1994 Positional cloning of the mouse obese gene and its human homologue. *Nature* **372** 425-432.

TABLE CAPTIONS

Table 1: Primer sequences. Primer names that start with 'q' indicate the primers used for qPCR.

Table 2: List of BLAST hits following comparison between zebrafish leptin-a (A) and leptin-b (B) to the non-redundant protein database (nr). BLAST hits are scored by an 'E-value', which applies statistical probability that the similarity between two sequences is based on stochastic events.

Table 3: Percentages for amino acid sequence identities between vertebrate leptin sequences.

FIGURE LEGENDS

Figure 1: cDNA and deduced amino acid sequence of the coding sequence of zebrafish leptin-a and leptin-b. Accession numbers are AM920658 and AM901009 respectively.

Figure 2: Multiple sequence alignment of zebrafish leptins, carp leptins and human leptin. Asterisk indicate amino acids that are conserved in all sequences, whereas colons and dots reflect decreasing levels of amino acid similarity. The four α-helices (A-D) were inferred from human leptin and are boxed in the alignment. The cysteine residues that form leptin's single disulphide bridge are shaded. Accession numbers: zebrafish leptin-a: AM920658, zebrafish leptin-b: AM901009, carp leptin-a-I: AJ868357, carp leptin-a-II: AJ868356, human leptin: P41159.

Figure 3: The gene structures of zebrafish leptins and mammalian leptins are conserved. Boxes represent coding exons only and are drawn to scale. Numbers inside the boxes reflect exon sizes in nucleotides. The intron phase is indicated with underlined numbers.

Figure 4: Protein models of the duplicate zebrafish leptins and human leptin. Zebrafish leptin-a (B), leptin-b (C) and human leptin (A) were modelled on the human leptin crystal structure and conform to the four α-helix bundle adopted by human leptin. In yellow the single disulphide bridge that stabilizes leptin's tertiary structure, in red the third cysteine of zebrafish leptin-b.

Figure 5: Phylogenetic tree of vertebrate leptins. Numbers at the branches reflect the confidence level as obtained by bootstrapping (1000 replications). Growth hormone (GH) and ciliary neurotrophic factor (CNTF) (both class-I helical cytokines) were included as outgroup. Only full length sequences were used for phylogenetic analysis. Accession numbers are as follows: chimpanzee leptin: O02750, human leptin: P41159, mouse leptin: P41160, dog leptin: O02720, cattle leptin: P50595, fat-tailed dunnart leptin:AF159713, Sout-African clawed frog leptin: AY884210, carp leptin-a-I: AJ836745, carp leptin-a-II: AJ836744, zebrafish leptin-a: AM920658, rainbow trout leptin: AB354909, zebrafish leptin-b: AM901009, medaka leptin-a: AB193548, medaka leptin-b: BN001183, tiger pufferfish leptin: AB193547, green-spotted pufferfish leptin: AB193549, human GH: P01241, zebrafish GH: Q1JQ34, human CNTF: P26441, mouse CNTF: P51642.

Figure 6: The synteny between the human leptin locus and both zebrafish leptin loci is conserved. A comparison between the human leptin locus (7q32.1) and the zebrafish leptin loci (located on chromosome 18 and 4 respectively) reveals that adjacent to both zebrafish leptins there are multiple genes that lie adjacent to human leptin. Arrows reflect genes, the direction of the arrow the orientation of the gene. Black arrows represent leptin orthologues, grey arrows represent genes in synteny in the human and zebrafish leptin loci. Genes are not drawn to scale, nor is intergenic space included. Abbreviations: ARF5: ADP-ribosylation factor 5, RBM28: RNA Binding Protein Motif 28, SND1: Staphylococcal nuclease domain-containing protein 1, GCC1: GRIP and coiled-coil domain containing 1.

Figure 7: Basal expression of leptin-a (open bars) and leptin-b (closed bars). Leptin-a and leptin-b are constitutively expressed in all organs investigated. Bars represent the mean value of four individual zebrafish. Error bars indicate standard errors. Note the logarithmic scale of the x-axis.

Figure 8: Leptin-a and leptin-b mRNA expression after one week fasted (closed bars) and fed (control; open bars) zebrafish. Leptin-b mRNA decreases significantly (*: $P<0.05$) after fasting for one week. Bars represent the mean value, error bars indicate standard errors.

Figure 9: General phylogenetic tree of vertebrate evolution. Mammals and teleosts shared their last common ancestor ~450 Mya. The finding of duplicate leptin paralogues in the medaka (*Beloniformes*) and zebrafish (*Cypriniformes*) dates the duplication event that gave rise to the duplicated leptins to ~296 Mya, as these species shared their last common ancestor at that time point. The tetraploidization of the carp genome (~16 Mya) is likely the event that gave rise to paralogous leptin-a-I and leptin-a-II genes in carp. Divergence estimates are based on: (Hedges 2002; Hoegg and Meyer 2005; Volff 2005; Zardoya and Doadrio 1999).

Gene	Accession Nr.	Primer	Sequence 5´→ 3´
leptin-a	AM920658	zf.leptin-a.fw	ATG CGT TTT CCA GCT CTC
		zf.leptin-a.rv	TCA GCA GAT TTT CAG CTG GTC
		Q-zf.leptin-a.fw	GAC TGC ACA CTG AAG GAA TC
		Q-zf.leptin.a.rv	GCA CTG TCC TCT AGA AAA GC
leptin-b	AM901009	zf.leptin-b.fw	ATG AAG TCT TCA ATG ATT TTT TGC
		zf.leptin-b.rv	CAG AGA ATG AAT GTC TCA GCC
		Q-zf.leptin-b.fw	ATT GCT CGA ACC ACC ATC AG
		Q-zf.leptin-b.rv	GAT GTC AGG GCC GAA ATC AA
40S ribosomal protein S11	CA472846	Q-40S.fw	AAA CAG CCC ACC ATC TTC CA
		Q-40S.rv	CTG TGA TAA CGA GGG AGC TTT TC
β-actin	AF025305	Q-BACT.fw	CAA CAG GGA AAA GAT GAC ACA GAT
		Q-BACT.rv	CAG CCT GGA TGG CAA CGT

A

Accession number	Species	Description	E-value
BN000380	Zebrafish	Leptin-a	$5*10^{-89}$
AJ868357	Common carp	Leptin-a-I	$7*10^{-49}$
AJ868356	Common carp	Leptin-a-II	$2*10^{-48}$
ABV57772	Goldfish	Leptin-a-II	$2*10^{-32}$
AAZ66785	Channel catfish	Leptin	$2*10^{-14}$
AY884210	*Xenopus laevis*	Leptin	$6*10^{-9}$
AAY68394	Tiger salamander	Leptin	$9*10^{-6}$
AM901009	Zebrafish	Leptin-b	$2*10^{-5}$

B

Accession number	Species	Description	E-value
AM901009	Zebrafish	Leptin-b	$7*10^{-58}$
AJ868356	Common carp	Leptin-a-II	$5*10^{-8}$
AJ868357	Common carp	Leptin-a-I	$2*10^{-7}$
ABV57772	Goldfish	Leptin-a-II	$9*10^{-7}$
BN000380	Zebrafish	Leptin-a	$9*10^{-7}$
AAZ66785	Channel catfish	Leptin	$1*10^{-4}$

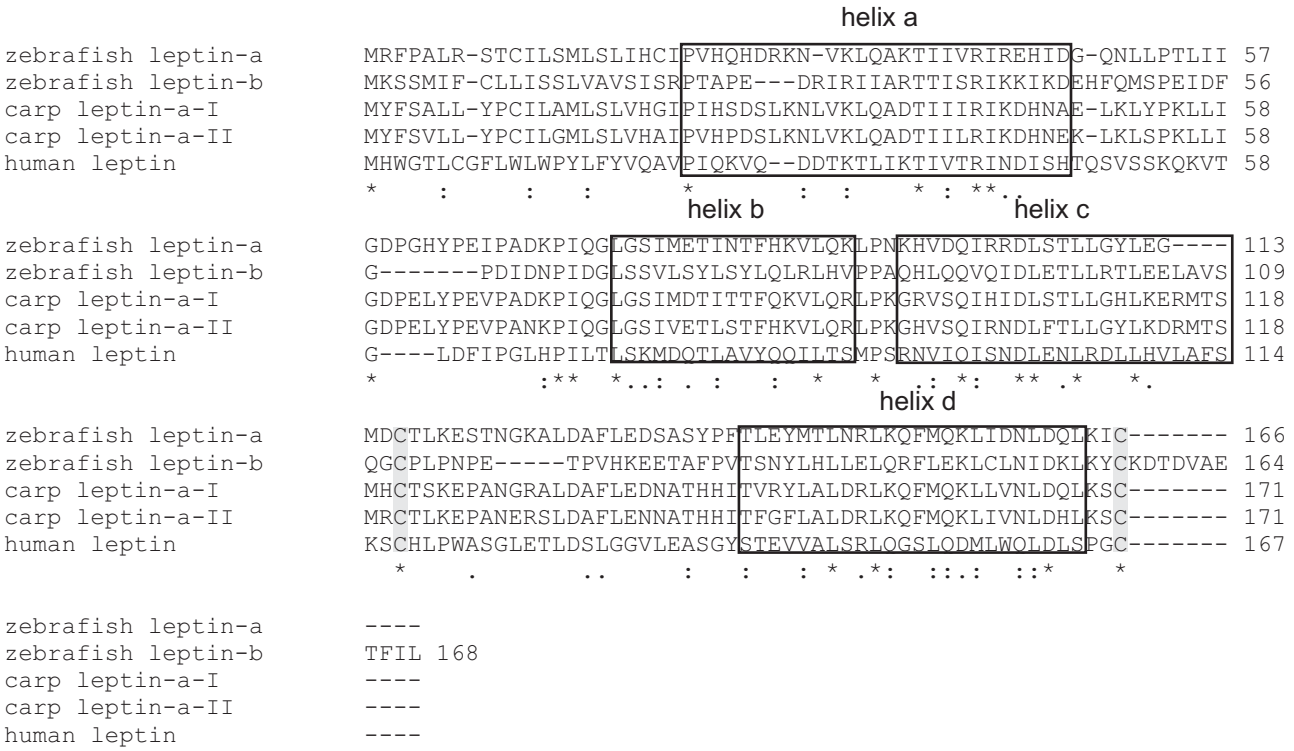
	zebrafish leptin-a	zebrafish leptin-b	carp leptin-a-I	carp leptin-a-II	clawed frog	human	mouse	dog	cow
zebrafish leptin-a	100								
zebrafish leptin-b	24	100							
carp leptin-a-I	64	25	100						
carp leptin-a-II	63	25	81	100					
clawed frog	24	13	27	27	100				
human	19	19	21	23	35	100			
mouse	21	19	23	25	34	83	100		
dog	19	18	22	23	33	80	78	100	
cow	20	19	23	25	34	84	83	88	100

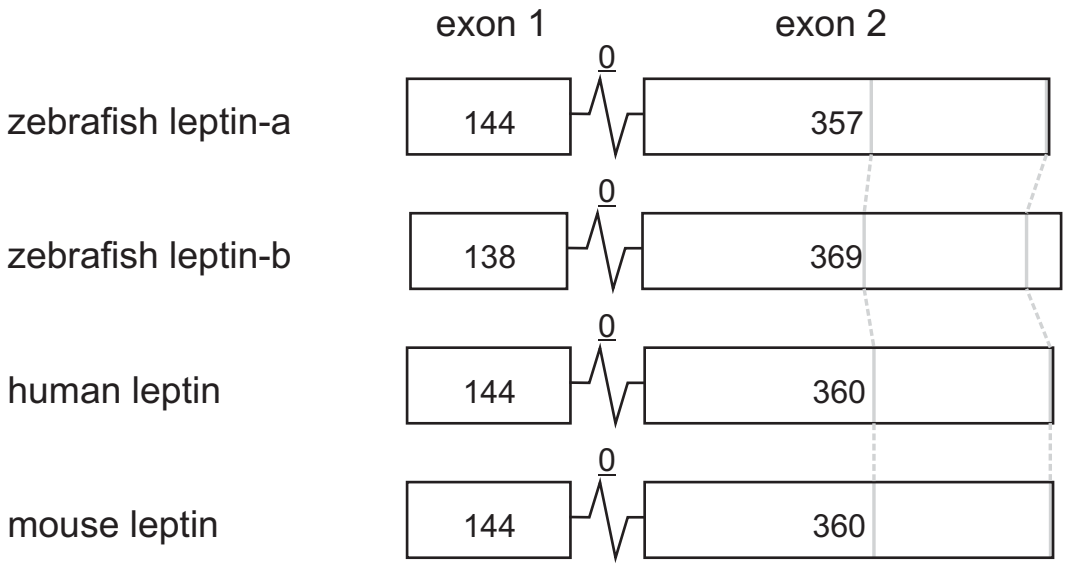
A

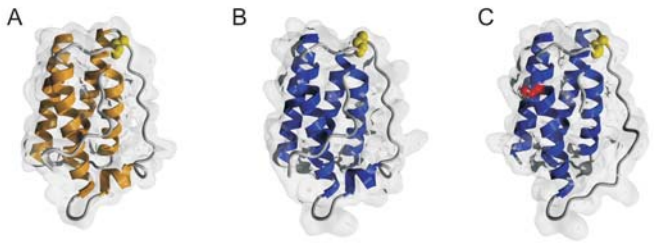
```
  1 M  R  F  P  A  L  R  S  T  C  I  L  S  M  L  S  L  I  H  C
  1 atgcgttttccagctctccgctcaacctgtattttgagcatgctcagtttgattcattgc
 21 I  P  V  H  Q  H  D  R  K  N  V  K  L  Q  A  K  T  I  I  V
 61 attcccgttcatcagcatgaccggaaaaatgtcaaactgcaggcaaagaccatcatcgtc
 41 R  I  R  E  H  I  D  G  Q  N  L  L  P  T  L  I  I  G  D  P
121 agaatcagggaacacattgacgggcaaaatttacttccaacgctcatcattggggatcca
 61 G  H  Y  P  E  I  P  A  D  K  P  I  Q  G  L  G  S  I  M  E
181 ggacattatccagagattcccgctgacaaacccatccaagggctcggctccatcatggaa
 81 T  I  N  T  F  H  K  V  L  Q  K  L  P  N  K  H  V  D  Q  I
241 accattaataccttccacaaggttcttcagaagcttccaaataagcatgttgaccagata
101 R  R  D  L  S  T  L  L  G  Y  L  E  G  M  D  C  T  L  K  E
301 cgccgagatctatccacacttctgggttacctggaaggcatggactgcacactgaaggaa
121 S  T  N  G  K  A  L  D  A  F  L  E  D  S  A  S  Y  P  F  T
361 tcaacaaatgggaaagcgctggacgcttttctagaggacagtgcttcatatccctttact
141 L  E  Y  M  T  L  N  R  L  K  Q  F  M  Q  K  L  I  D  N  L
421 ttagagtacatgactttaaacagactgaaacagtttatgcaaaagctgatcgataatctg
161 D  Q  L  K  I  C  *
481 gaccagctgaaaatctgctga
```

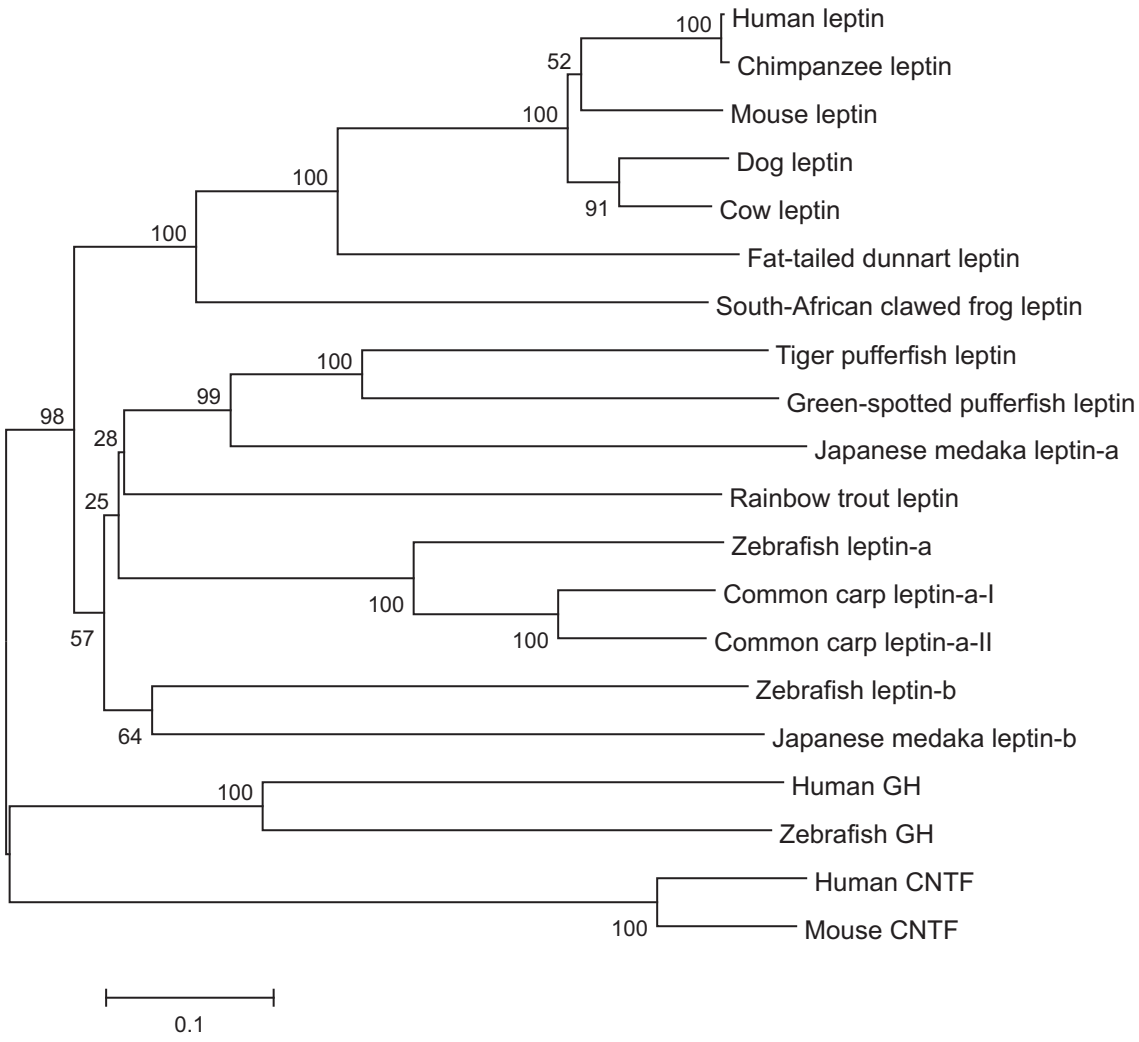
B

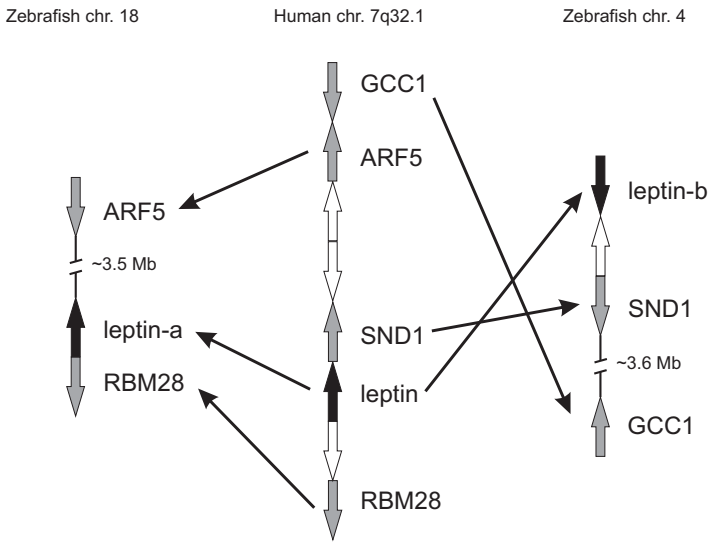
```
  1 M  K  S  S  M  I  F  C  L  L  I  S  S  L  V  A  V  S  I  S
  1 atgaagtcttcaatgatttttgcttgttaatatcatccctggtggccgtgagcatcagt
 21 R  P  T  A  P  E  D  R  I  R  I  I  A  R  T  T  I  S  R  I
 61 cgacccacggctcccgaagacaggatacgaatcattgctcgaaccaccatcagccgaatt
 41 K  K  I  K  D  E  H  F  Q  M  S  P  E  I  D  F  G  P  D  I
121 aaaaaaatcaaagatgagcacttccagatgtctccagagattgatttcggccctgacatc
 61 D  N  P  I  D  G  L  S  S  V  L  S  Y  L  S  Y  L  Q  L  R
181 gacaaccccattgatggtctcagttctgtcttgagttacttgagttacctgcagttgcgg
 81 L  H  V  P  P  A  Q  H  L  Q  Q  V  Q  I  D  L  E  T  L  L
241 ttgcatgttcctccagctcagcacctacagcaggtccagatagacttagagactctcctg
101 R  T  L  E  E  L  A  V  S  Q  G  C  P  L  P  N  P  E  T  P
301 aggacactggaggaactggccgtctcacagggatgccctctacccaatcccgagaccccg
121 V  H  K  E  E  T  A  F  P  V  T  S  N  Y  L  H  L  L  E  L
361 gtgcataaagaagaaacagccttccccgtcacctccaactacctgcacctcctggagctc
141 Q  R  F  L  E  K  L  C  L  N  I  D  K  L  K  Y  C  K  D  T
421 cagaggttcctggagaagctctgcctcaacatagacaaactgaaatactgcaaagataca
161 D  V  A  E  T  F  I  L  *
481 gatgtggctgagacattcattctctga
```

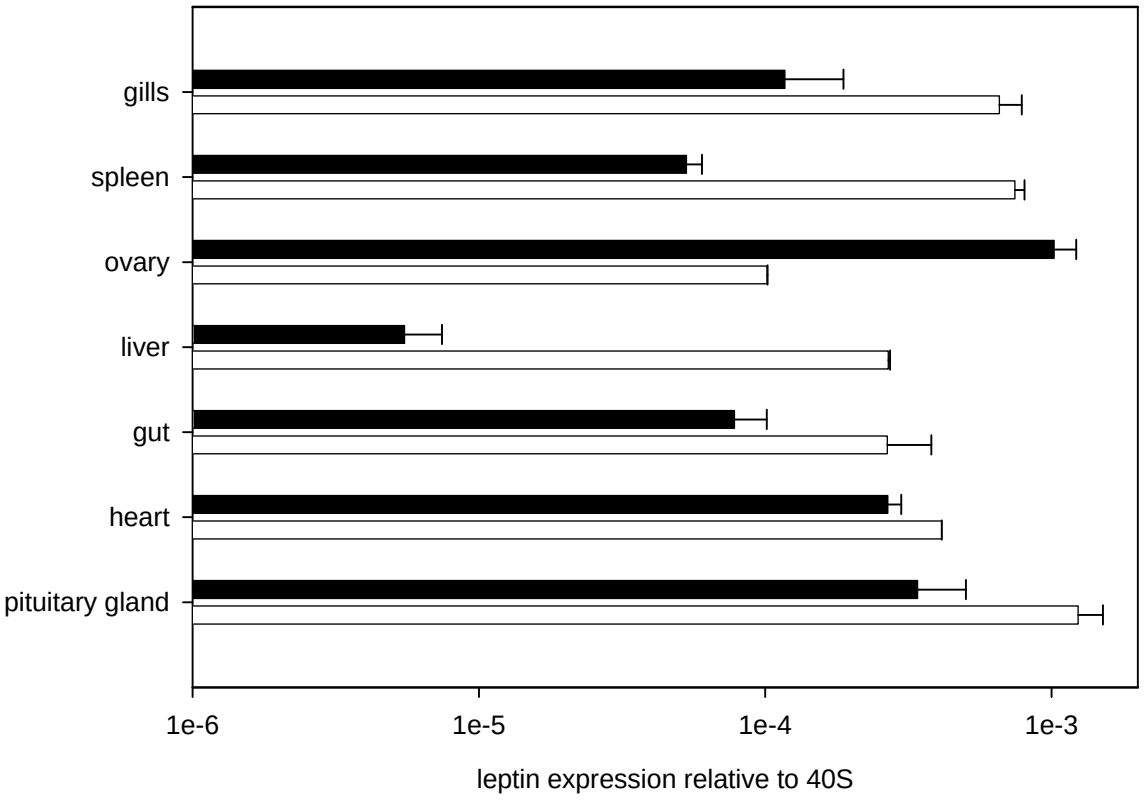

helix a
helix b
helix c
helix d
zebrafish leptin-a MRFPALR-STCILSMLSLIHCIPVHQHDRKN-VKLQAKTIIVRIREHIDG-QNLLPTLII 57
zebrafish leptin-b MKSSMIF-CLLISSLVAVSISRPTAPE---DRIRIIARTTISRIKKIKDEHFQMSPEIDF 56
carp leptin-a-I MYFSALL-YPCILAMLSLVHGIPIHSDSLKNLVKLQADTIIIRIKDHNAE-LKLYPKLLI 58
carp leptin-a-II MYFSVLL-YPCILGMLSLVHAIPVHPDSLKNLVKLQADTIILRIKDHNEK-LKLSPKLLI 58
human leptin MHWGTLCGFLWLWPYLFYVQAVPIQKVQ--DDTKTLIKTIVTRINDISHTQSVSSKQKVT 58
zebrafish leptin-a GDPGHYPEIPADKPIQGLGSIMETINTFHKVLQKLPNKHVDQIRRDLSTLLGYLEG---- 113
zebrafish leptin-b G-------PDIDNPIDGLSSVLSYLSYLQLRLHVPPAQHLQQVQIDLETLLRTLEELAVS 109
carp leptin-a-I GDPELYPEVPADKPIQGLGSIMDTITTFQKVLQRLPKGRVSQIHIDLSTLLGHLKERMTS 118
carp leptin-a-II GDPELYPEVPANKPIQGLGSIVETLSTFHKVLQRLPKGHVSQIRNDLFTLLGYLKDRMTS 118
human leptin G----LDFIPGLHPILTLSKMDQTLAVYQQILTSMPSRNVIQISNDLENLRDLLHVLAFS 114
zebrafish leptin-a MDCTLKESTNGKALDAFLEDSASYPFTLEYMTLNRLKQFMQKLIDNLDQLKIC------- 166
zebrafish leptin-b QGCPLPNPE-----TPVHKEETAFPVTSNYLHLLELQRFLEKLCLNIDKLKYCKDTDVAE 164
carp leptin-a-I MHCTSKEPANGRALDAFLEDNATHHITVRYLALDRLKQFMQKLLVNLDQLKSC------- 171
carp leptin-a-II MRCTLKEPANERSLDAFLENNATHHITFGFLALDRLKQFMQKLIVNLDHLKSC------- 171
human leptin KSCHLPWASGLETLDSLGGVLEASGYSTEVVALSRLQGSLQDMLWQLDLSPGC------- 167
zebrafish leptin-a ----
zebrafish leptin-b TFIL 168
carp leptin-a-I ----
carp leptin-a-II ----
human leptin ----

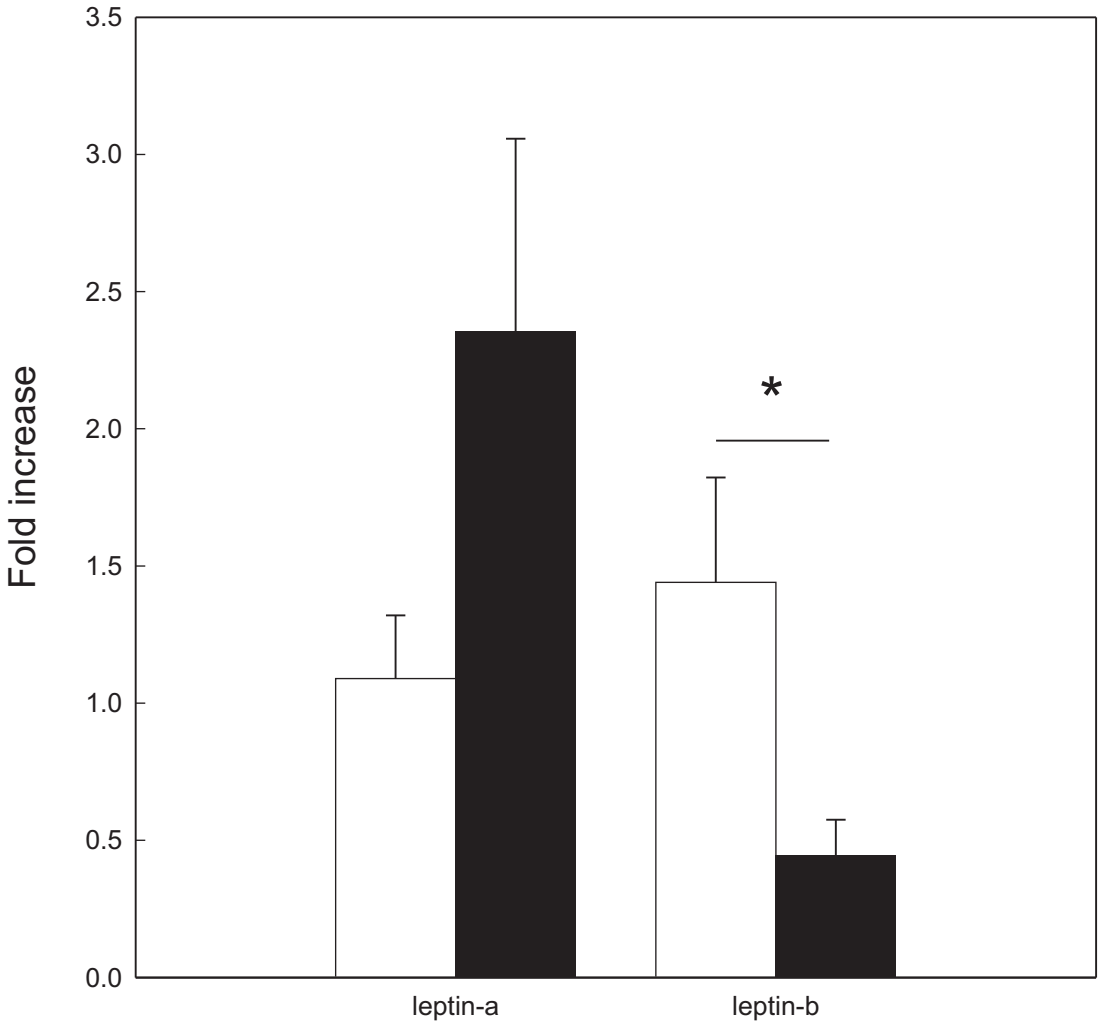

exon 1
exon 2
zebrafish leptin-a
144
0
357
zebrafish leptin-b
138
0
369
human leptin
144
0
360
mouse leptin
144
0
360

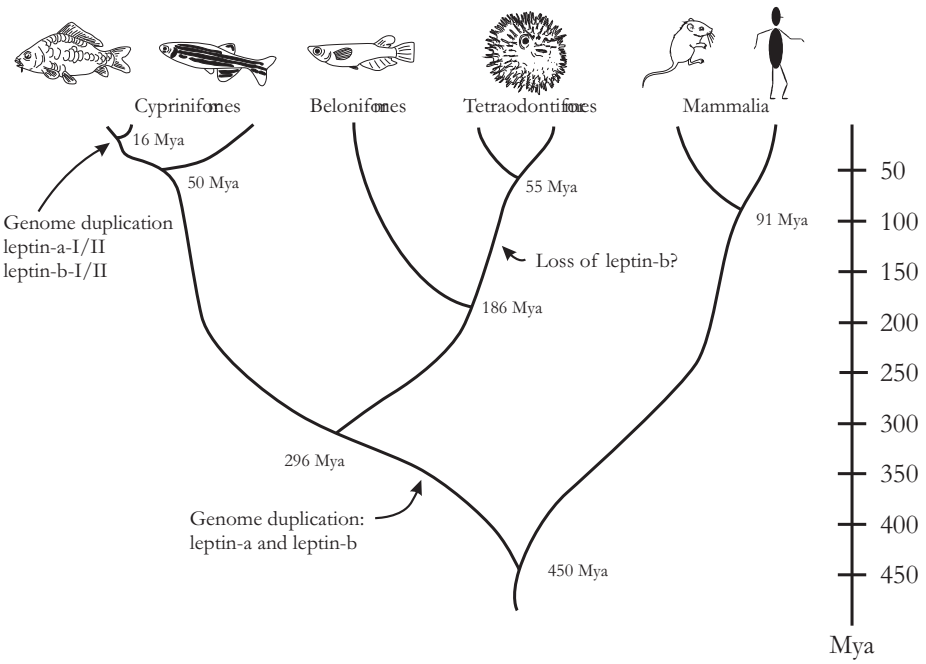

A
B
C


Human leptin
Chimpanzee leptin
Mouse leptin
Dog leptin
Cow leptin
Fat-tailed dunnart leptin
South-African clawed frog leptin
Tiger pufferfish leptin
Green-spotted pufferfish leptin
Japanese medaka leptin-a
Rainbow trout leptin
Zebrafish leptin-a
Common carp leptin-a-I
Common carp leptin-a-II
Zebrafish leptin-b
Japanese medaka leptin-b
Human GH
Zebrafish GH
Human CNTF
Mouse CNTF
100
52
100
100
91
100
100
99
98
28
25
100
100
57
64
100
100
0.1


Zebrafish chr. 18
Human chr. 7q32.1
Zebrafish chr. 4
GCC1
ARF5
SND1
leptin
RBM28
ARF5
~3.5 Mb
leptin-a
RBM28
leptin-b
SND1
~3.6 Mb
GCC1


gills
spleen
ovary
liver
gut
heart
pituitary gland
1e-6
1e-5
1e-4
1e-3
leptin expression relative to 40S


Fold increase
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
*
leptin-a
leptin-b


Cypriniformes
Beloniformes
Tetraodontiformes
Mammalia
16 Mya
50 Mya
55 Mya
91 Mya
186 Mya
296 Mya
450 Mya
Genome duplication
leptin-a-I/II
leptin-b-I/II
Loss of leptin-b?
Genome duplication:
leptin-a and leptin-b
50
100
150
200
250
300
350
400
450
Mya