

Multiple transcripts encoding lamprey gonadotropin-releasing hormone-I precursors

K Suzuki, R L Gamble and S A Sower

Department of Biochemistry and Molecular Biology, Rudman Hall, University of New Hampshire, Durham, New Hampshire 03824, USA

(Requests for offprints should be addressed to S A Sower; Email: sasower@cisunix.unh.edu)

ABSTRACT

The cDNA encoding lamprey prepro-gonadotropin releasing hormone-I (lamprey GnRH-I) has been isolated and sequenced in an agnathan, the sea lamprey, *Petromyzon marinus*. The lamprey GnRH-I precursor is the first identified in an ancient lineage of vertebrates and has the same overall tripartite structure as other vertebrate GnRH precursors. The amino acid sequence of lamprey GnRH-I and the processing site (Gly-Lys-Arg) are highly conserved during 500 million years of evolution with 60–70% identity compared with those of tetrapod and teleost GnRH precursors. In contrast, the GnRH associated peptide regions are markedly divergent, with less than 20% identity compared with all identified vertebrate precursors.

Unlike all other known vertebrate GnRH precursors, which typically have one and in a single case two transcripts, three distinct transcripts were isolated and sequenced in lampreys. These lamprey GnRH-I transcripts, termed GAP49, GAP50 and GAP58, differed in the length of the GAP coding sequence and were demonstrated to be the products of a single gene. Analysis of the lamprey GnRH-I gene intron-2 splice junction demonstrated that alternate splicing produces the different lamprey GnRH-I transcripts. Lamprey GnRH-I is the first GnRH gene demonstrated to utilize splice sequence variants to produce multiple transcripts, which may reflect an ancestral gene regulatory mechanism. *Journal of Molecular Endocrinology* (2000) **24**, 365–376

INTRODUCTION

Gonadotropin-releasing hormone (GnRH) is the key regulatory neurohormone central to the control of reproduction in all vertebrates. Considerable diversity exists in the molecular structure of GnRHs among vertebrates and protochordates. Currently, ten structures of the GnRH decapeptide have been determined in various vertebrate species, including lamprey GnRH-I and -III, and two in invertebrates (Sower *et al.* 1993, 1998, Sherwood *et al.* 1986). In all GnRH peptides, certain regions of the molecule have been highly conserved, including the NH₂-terminal, pGlu¹ and Ser⁴, and the COOH-terminal. These regions and the length of the molecule have remained unchanged during 500 million years of chordate evolution.

Similar to other neuropeptides, GnRH is first synthesized as a larger precursor protein, prepro-GnRH, which is then processed to the final decapeptide (Klungland *et al.* 1992a). The nucleotide sequences of GnRH have been isolated from human (Seeburg & Adelman 1984, Adelman *et al.*

1986, Hayflick *et al.* 1989, Radovick *et al.* 1990, Dong *et al.* 1993, White *et al.* 1998), rat (Adelman *et al.* 1986, Goubau *et al.* 1992), mouse (Mason *et al.* 1986), frog (Hayes *et al.* 1994), salmonid species (Coe *et al.* 1995, Klungland *et al.* 1992a, b, Suzuki *et al.* 1992, Ashihara *et al.* 1995), African cichlid (Bond *et al.* 1991, White *et al.* 1994, 1995), red seabream (Okuzawa *et al.* 1994), plainfin midshipman (Grober *et al.* 1995), striped bass (Gothilf *et al.* 1995, Chow *et al.* 1998), gilthead seabream (Gothilf *et al.* 1996), goldfish (Lin & Peter 1996), catfish (Bogerd *et al.* 1994) and chicken (Dunn *et al.* 1993). Although the nucleotide sequences of six forms of GnRH have been isolated and sequenced from different species representing four of the seven classes of vertebrates, to date a GnRH nucleotide sequence has not been identified in any species from the two oldest lineages of vertebrates, the Agnatha and Chondrichthyes, or from the invertebrates.

From the known nucleotide sequences of GnRH, it is clear that the precursor molecule is tripartite. It consists of a signal peptide, the GnRH decapeptide

directly followed by a cleavage site, and the sequence encoding a GnRH-associated peptide (GAP) (Andersen & Klungland 1993). The genomic sequences for GnRH have been determined in human and rat (Adelman *et al.* 1986), Atlantic salmon (Klungland *et al.* 1992b), chicken (Chicken GnRH-I; Dunn *et al.* 1993), masu salmon (Higa *et al.* 1997), striped bass (Chow *et al.* 1998), and sockeye salmon (Coe *et al.* 1995). The organization of the gene is similar across these species consisting of four exons. The coding region is dispersed over exons 2–4. Exon-2 encodes the signal peptide, the GnRH decapeptide, the cleavage site and the N-terminal portion of GAP; exon-3 consists entirely of the middle portion of the GAP molecule; exon-4 contains the C-terminal of GAP and the 3' untranslated region (Andersen & Klungland 1993).

The objective of the present study was to isolate and characterize the cDNA that encodes lamprey GnRH. This information will allow for a more complete comparison of the evolution of genetic information encoding a significant regulatory peptide key in the reproductive control of vertebrate species.

MATERIALS AND METHODS

Animals

Adult sea lampreys were captured during their upstream migration at the Cochecho River salmon ladder in Dover, NH, USA. The lampreys were maintained at the Anadromous Fish and Invertebrate Research Laboratory at the University of New Hampshire.

Nucleic acid preparation

Total RNA was prepared typically from 1 g frozen tissue using the single step method described by Chomzynski & Sacchi (1987). The integrity of total RNA was verified by ethidium bromide staining of 28S and 18S ribosomal RNA. Poly(A⁺) RNA was affinity purified using either a Poly(A⁺) RNA Quick isolation kit (Stratagene, La Jolla, CA, USA) or by a PolyA tract mRNA isolation system (Promega, Madison, WI, USA). First-strand cDNA for PCR template was synthesized using the avian myeloblastosis virus reverse transcriptase (AMV-RT) and oligo (dT)₂₅ as a primer.

Isolation of lamprey GnRH-I cDNA

The lamprey GnRH-I cDNA was obtained by rapid amplification of cDNA ends (RACE)

(Frohman *et al.* 1988). A degenerate primer, L-GnRH32 (5'-AGCA(C/T)T(G/A)ITCI(C/T)A(C/T)(G/C)AITGGAA(G/A)CCIGGGIGG-3') was designed according to the peptide sequence of lamprey GnRH-I and -III. PCR amplification of the 3' end of cDNA was performed using L-GnRH32 and oligo (dT)₂₅ primer. The first 15 cycles of PCR amplification were carried out under the following conditions: denaturation at 94 °C for 30 s, annealing at 40 °C for 1 min, extension at 72 °C for 1 min. The subsequent 20 cycles of amplification were performed with the following conditions: denaturation at 94 °C for 45 s, annealing at 60 °C for 30 s, extension at 72 °C for 45 s. PCR fragments were cloned into a pT7-Blue vector (Novagen, Madison, WI, USA), T-easy vector (Promega) or p-GEM 7 (Promega) for sequence analysis by the University of New Hampshire sequencing facility.

Next, an antisense primer specific to the putative GnRH GAP region was designed for 5'-RACE. The 5' RACE of lamprey GnRH-I cDNA was performed using the antisense GAP primer and poly (dG)₁₅. The forty cycles of PCR amplification were carried out as follows: denaturation at 94 °C for 45 s, annealing at 40 °C for 1 min, extension at 72 °C for 1 min. The PCR fragments were subcloned and subjected to Southern blot analysis to select clones containing lamprey GnRH-I sequence.

Specific PCR primers (RG1, RG2, RG3 and RG4) were designed according to the partial lamprey GnRH-I cDNA sequences obtained by RACE. The primers were used in the combinations RG1/RG4, RG1/RG2 and RG3/RG4 to sequence lamprey GnRH-I completely. The PCR fragments were purified by Gene Clean II (BIO 101 Inc.) and either sequenced directly or cloned into plasmid vectors for sequencing.

Northern blot and RT-PCR analysis

The tissue distribution of lamprey GnRH-I mRNA was examined by both northern blot and reverse transcriptase PCR (RT-PCR) analyses. For northern blot analysis 10 µg mRNA from lamprey brain, liver, kidney, testis and ovary were separated on a 1.5% agarose 6% formaldehyde gel and transferred to a Hybond-N nylon membrane (Amersham). Lamprey GnRH-I mRNA was detected by the hybridization (6 × SSPE, 0.1% SDS, 5 × Denhardt's solution; 67 °C, 16 h) of radiolabeled lamprey GnRH-I cDNA, which was prepared by PCR amplification of the cDNA using RG-1 and RG-4 as primers.

For RT-PCR, PCR products of mRNA isolated from lamprey brain, liver, kidney, testis and ovary

were amplified using the GnRH-I specific primer pairs RG1/RG2 and RG1/RG4 in 35 cycles of the following conditions: denaturation at 94 °C for 30 s, annealing at 50 °C for 1 min, extension at 72 °C for 2 min. PCR fragments were separated by 1% agarose gel electrophoresis and transferred to nylon membranes for Southern blot analysis (Sambrook *et al.* 1989).

Lamprey GnRH-I cDNA (obtained from 3'-RACE) was used as the probe to determine the authenticity of PCR fragments as lamprey GnRH-I. As a control, β -actin cDNA was amplified from the same tissues and analyzed using mouse β -actin cDNA as the probe (Leader *et al.* 1986).

Genomic DNA analysis

To determine gene copy number of lamprey GnRH-I, lamprey genomic DNA was digested with the restriction enzymes, Apa-I (Promega), Bstx-I (Promega), EcoR-I (Stratagene) and Pst-I (Stratagene). Digested DNA was separated by agarose gel electrophoresis and transferred to a nylon membrane for Southern blot analysis using the lamprey GnRH-I cDNA as a probe. Lamprey genomic DNA was further analyzed in order to examine the intron-2 and intron-3 splice junctions. Genomic DNA was extracted from lamprey testis as described by Sambrook *et al.* (1989). Two sets of primers, intron-2⁺/intron-2⁻ and intron-3⁺/RG-2, were used for PCR amplification of genomic DNA in 35 cycles of the following conditions: denaturation at 94 °C for 30 s, annealing at 50 °C for 1 min, extension at 72 °C for 2 min. PCR fragments separated by electrophoresis were extracted from agarose gels by the freeze/squeeze procedure (Tautz & Renz 1983), and cloned into a T-easy vector (Promega) for sequence analysis.

Sequence and phylogenetic analyses

The signal peptide cleavage site and the transmembrane helix structure of lamprey GnRH-I were predicted by the methods described by Nielsen *et al.* (1997) and Sonnhammer *et al.* (1998) respectively. The sequence homology of prepro-GnRH was analyzed by Megalign (DNA Star) using the PAM250 weight table for multiple alignment of prepro-GnRH sequences. Each family (or group) of GnRHs was aligned separately with the lamprey GnRH precursor to prevent gap insertions. Hydrophilicity of prepro-GnRH was estimated by the Kyte–Doolittle method using the program Protean in DNA Star (Kyte & Doolittle 1982). The α -helix regions of the GnRH precursor

were predicted using the Garnier–Robson method (Garnier & Robson 1989).

Thirty-two GnRH prehormones or 24 cDNA sequences encoding prepro-GnRH were used for phylogenetic analysis of the lamprey GnRH-I precursor using Megalign. Species and GenBank accession numbers of GnRH sequences used are as follows: pig L32864; mouse M14972; rat M15527; sheep U02517; tree shrew U63326 & U63327; human E00844 & AF036329; chicken X69491; *Xenopus* L28040; roach U60667 & U60668; goldfish U30386 & U30301; sockeye salmon D31869 & D31868; cherry salmon D10946; Atlantic salmon X79709; brook trout X79712; plainfin midshipman U41669; red seabream D86582; gilthead sea bream U30311, U30320 & U30325; cichlid AF076963, AF076962 & AF076961; African catfish X78049 & X78047; striped bass AF056314; guinea pig AF033346. The phylogenetic tree was constructed by PAUP 4.0 (Swofford 1998) using the distances between each prepro-GnRH sequence as parameters.

RESULTS

Isolation of lamprey GnRH-I cDNA

The 3'-RACE amplified the region of lamprey GnRH-I cDNA from the GnRH decapeptide to the poly(A) tail (Fig. 1). Sequence analysis of L-GnRH-3' identified partial lamprey GnRH-I sequences with a dibasic cleavage site (-Lys-Arg-) and a putative associated peptide sequence following the L-GnRH25 primer sequence. Clone L-GnRH-5', obtained by 5' RACE, also contained partial lamprey GnRH-I sequence. L-GnRH-3' and L-GnRH-5' sequences overlapped by 133 bp. However, additional clones representing both the 3'- and 5'-ends of lamprey GnRH-I cDNA were analyzed. The entire cDNA sequence for lamprey GnRH-I was obtained by overlapping the sequences of the PCR clones created with the primer pairs RG1/RG4, RG1/RG2 and RG3/RG4.

As shown in Fig. 2, the lamprey GnRH-I cDNA consists of 638 nucleotides. The deduced amino acid (aa) sequence of prepro-lamprey GnRH-I is 87 residues in length. Like other known prepro-GnRHs, the lamprey GnRH-I precursor is tripartite and consists of a signal peptide (24 aa), a GnRH peptide with a processing site (13 aa) and a GAP (50 aa). The predicted signal peptide cleavage site is located just before the first glutamine of the GnRH decapeptide. Transmembrane helix structure was found at the amino-terminal part of the lamprey GnRH-I precursor, which is located between serine

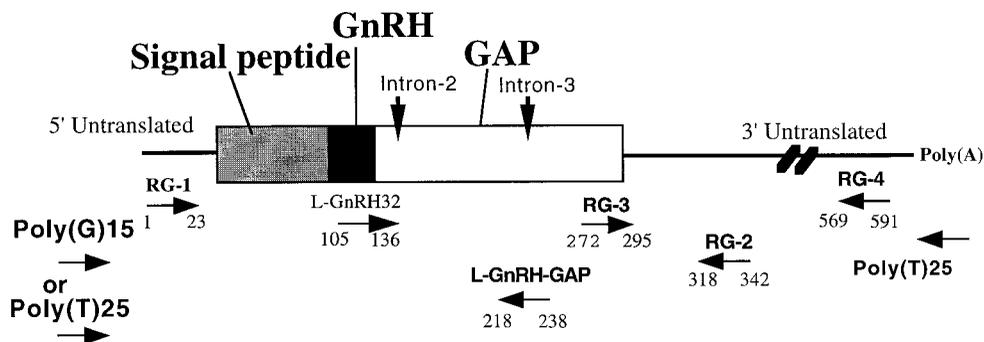


FIGURE 1. The corresponding primer positions in lamprey GnRH-I cDNA and the PCR products obtained by 3'- and 5'-RACE. Primer positions used for PCR amplification (RACE and subsequent RT-PCR) and for sequencing of lamprey GnRH-I cDNA. Arrows represent the direction of the primer and the number under the arrow represents the primer position corresponding to the lamprey GnRH-I cDNA.

(-19) of the signal peptide and leucine (+5) of the GnRH decapeptide.

Multiple lamprey GnRH-I transcripts

Heterogeneity was detected in the sequences coding for GAP. In 25 clones that were analyzed,

three different transcripts were identified: lamprey GnRH-I GAP49, lamprey GnRH-I GAP50 and lamprey GnRH-I GAP 58 (GenBank Accession Numbers: AF144481, AF144479, AF144480 respectively). The DNA sequences encoding the untranslated regions, the signal peptide and GnRH

tgctgtcgagagaagaacccttggaagacATGGCACTGCGCGGTGAGAGTCTGACACTCTGCTGCTGGCGACGGCTCTGCTCGTC 89
 MetAlaLeuArgGlyGlnSerLeuThrLeuLeuLeuLeuAlaThrAlaLeuLeuVal -6
 -24 **Signal peptide**

TCCTTGAATTACGGCGAGCACTACTCCCTGGAAATGGAAACCCGGGGGCAAACGAGACCTGGAGCAGGAGCTTGAGCCGCCGAGCAACGGC 179
 SerLeuAsnTyrAlaGlnHisTyrSerLeuGluTrpLysProGlyGlyLysArgAspLeuGluGlnGluLeuGluProProSerAsnAla +25
 +1 **L-GnRH-I Processing site**

TTCGAGTGCACGGACCCGAATGCGCCTTCTCTCGAGTGCCCAACACCAAGCTGATCAGGGAGCTCGCGAGCTACCTCTCGCAGAGGAAT 269
 PheGluCysAspGlyProGluCysAlaPheSerArgValProAsnThrLysLeuIleArgGluLeuAlaSerTyrLeuSerGlnArgAsn +55
GnRH associated peptide

JAIGATCGGAAGGAGGCTGGAAGtagaccaccacgcatctcagggtgcagatgaagacgaccaaggtccgacacttcgcaacattcctc 359
 TyrAspArgLysGlyAlaLeuLysend +63
 ttaagcacgatgttagaacccttcaatgaacgcccttcaccttcgtcttaaccacgtgtattgcccgtgacgtctcctcgacacgt 449
 tgtgtccttaattacattcatcgccgctcagcgtatcgtaacatgatgtgaccgtacacatctttattgtgcagtacatagtagtactatg 539
 tttcagactggctgtcctgtctgtgggtgggatacactgtgctatgcgtggccatgtataacaaggaggtgtgtgtcataataaaattg 629

FIGURE 2. Nucleotide and deduced amino acid sequence of lamprey GnRH-I cDNA. Amino acid sequences with negative numbers correspond to the signal peptide of lamprey GnRH-I (underlined by a bold dotted line). The lamprey GnRH-I decapeptide sequence is enclosed in a square. The processing site is underlined by a bold solid line. GAP sequence is underlined by a dotted line. The location of polyadenylation signal (aataaa) is indicated by an arrow.

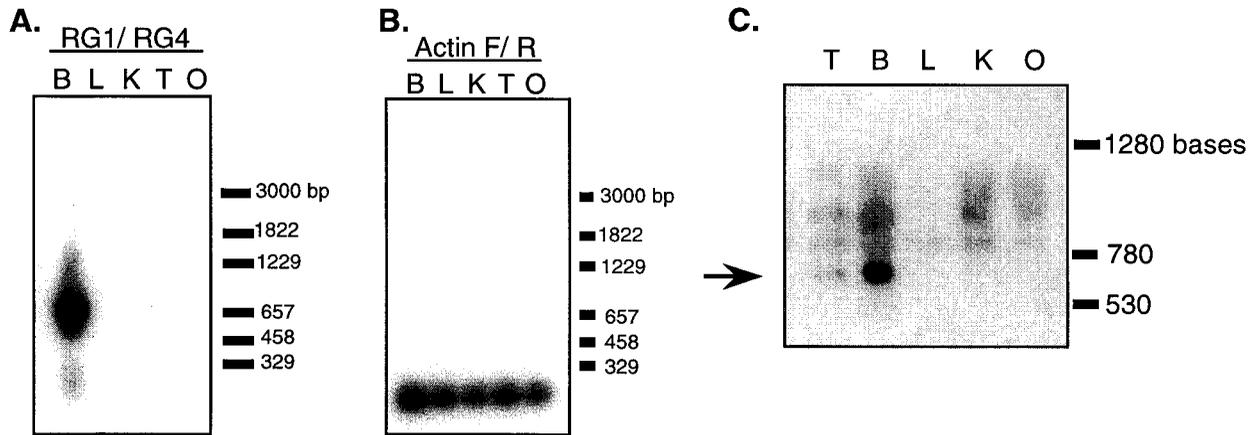


FIGURE 3. Expression of lamprey GnRH-I gene within various organs of mature sea lamprey (*P. marinus*). The amount of lamprey GnRH-I mRNA was estimated by RT-PCR and northern blot analysis. Expression of β -actin within various organs was also determined as control. B, Brain; L, liver; K, kidney; T, testis; O, ovary. (A) PCR fragments obtained with RG1/RG4 primers were analyzed by Southern blot using lamprey GnRH-I cDNA as probe. (B) As control, the cDNA of β -actin was amplified using mouse β -actin primers. PCR fragments were separated by gel electrophoresis and transferred to a nylon membrane. Radiolabeled mouse β -actin cDNA was used as probe. (C) Expression of lamprey GnRH-I in various organs was examined by northern blot analysis using lamprey GnRH-I cDNA as a probe. An arrow indicates the band of lamprey GnRH-I mRNA with an approximate size of 530 bases. A faint signal at the arrow was observed in testis RNA.

decapeptide were identical in all clones; only the GAP sequences varied. Nineteen clones of the 25 were lamprey GnRH GAP50, three clones were lamprey GnRH GAP49, and three were lamprey GnRH GAP58. The results from PCR analysis on cDNA prepared from a single lamprey brain confirmed that in the ten clones analyzed, one was lamprey GnRH GAP49, one was lamprey GnRH GAP58 and the remainder were lamprey GnRH-I GAP50. The DNA sequence of the predominant lamprey GnRH GAP50 was the original lamprey GnRH-I sequence obtained by direct sequencing of PCR fragments.

Tissue distribution of lamprey GnRH-I mRNA

Analysis by both RT-PCR (Fig. 3A) and northern blot analysis (Fig. 3C) revealed that the lamprey GnRH-I mRNA was primarily expressed in the brain. Preliminary results indicate minimal expression in the testis, but none in the liver, kidney and ovary.

Analysis of lamprey GnRH-I genomic DNA

All genomic DNA fragments encoding lamprey GnRH-I migrated as a single band, indicating that lamprey GnRH-I is encoded by a single gene (data not shown). The three lamprey GnRH-I transcripts

were thus believed to be the products of alternative splicing from a single gene. To explore the gene structure that produced three different lamprey GnRH-I transcripts, the genomic DNA fragments that contained complete intron-2 or intron-3 with adjacent exon sequences were amplified by PCR from lamprey genomic DNA (Fig. 4). Sequence analysis revealed several repeats of a putative splice site, GTCAGT, at the beginning of intron-2 (Fig. 5A). A tandem CAG sequence was also found at the end of intron-2 where the splice site for exon-3 was located. In contrast, a single splice site, GTGAGT, was found at the beginning and end of intron-3 (data not shown). Figures 5B and 6 present a model for the alternate splicing and GAP rearrangement of the three different lamprey GnRH-I transcripts.

Structural analysis of the lamprey GnRH-I precursor

Comparisons of the known prepro-GnRH sequences showed the region of the GnRH precursor with the highest similarity was the bioactive portion of the molecule – the GnRH decapeptide and the processing site (Fig. 7). Similarly to all other prepro-GnRHs, the lamprey GnRH-I precursor had a hydrophobic core and an α -helix structure located in the signal peptide.

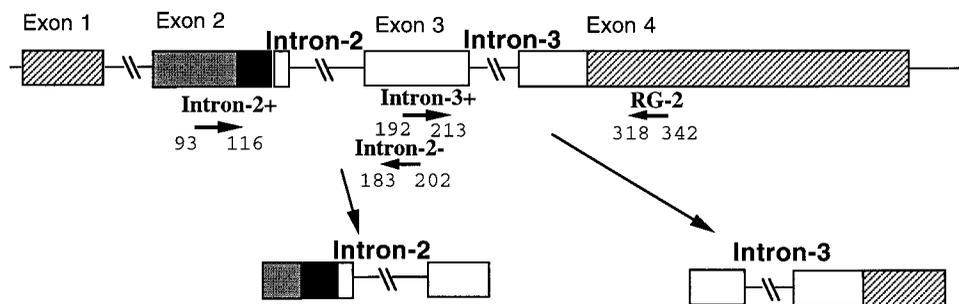


FIGURE 4. Analysis of genomic DNA coding lamprey GnRH-I. Location of primers used for the amplification of lamprey genomic DNA. The numbers under the arrows represent the corresponding positions of primers in the lamprey GnRH-I cDNA.

Phylogenetic analysis of GnRH precursors

The phylogenetic trees produced were based solely on the percent sequence dissimilarity between the precursors. Both the coding sequences and amino acid sequence of prepro-GnRHs were used to construct phylogenetic trees. In Fig. 8, only the tree for the coding sequences of GnRH cDNA is shown. The horizontal distance of each branch point indicates the relative dissimilarity between precursors shown in the tree.

DISCUSSION

The lamprey GnRH-I precursor is the first to be identified in an ancient lineage of vertebrates, from either the Agnatha or the Chondrichthyes. Novel to the GnRH family, three lamprey GnRH-I transcripts were isolated from lamprey brains. These transcripts differed in the length of the GAP coding sequence and were demonstrated to be the products of a single gene. Although the significance of three GnRH transcripts is unknown at this time, these analyses of the lamprey GnRH-precursor provide important information on the evolution and regulation of GnRH in vertebrates.

The lamprey GnRH-I decapeptide and the following processing site were between 60 and 70% identical to those of tetrapod and teleost GnRH precursors. This region of cDNA shared the greatest identity (72%) with that of chicken GnRH-II of striped bass (Chow *et al.* 1998). The lowest identity (53%) was found with guinea pig GnRH cDNA (Jimenez-Linan *et al.* 1997). Note that there are a number of different cDNAs of chicken GnRH-II that have been determined in different species of vertebrates. The processing site, Gly-Lys-Arg, has been well preserved throughout 500 million years of GnRH evolution. The least

conserved region (an average of 16% of homology) throughout evolution was GAP, but the common feature of GAP in GnRH precursors is that they are rich in hydrophilic amino acids and this feature is evident in the lamprey also. The average identity of GAP58 with other known GAPs was 14%. The GAP58 showed greatest identity (only 20%) with that of chicken GnRH-II of gilthead seabream. The lowest identity (11%) with GAP58 was observed in the GAP sequences of salmon GnRHs of cichlid.

The function of GAP has been examined in several studies, yet no definitive functions have been established. Human GAP was found to inhibit prolactin secretion in rat and GH3 cells (Nikolics *et al.* 1985, Van Chuoi *et al.* 1993) and was demonstrated to act through the inhibitory GTP-protein, Gi (Planas *et al.* 1990). The prolactin-inhibiting action of human GAP was also observed in the tilapia pituitary gland (Planas *et al.* 1990). However, this prolactin-inhibiting action was not demonstrated in other animals such as sheep (Thomas *et al.* 1988). A recent study of human GAP revealed that it has a helix-loop-helix structure that is necessary for prolactin inhibition, one of the proposed functions of GAP (Chavali *et al.* 1997). The lamprey GnRH-I GAP sequence does not share this secondary structure of human GAP. Instead, five short α -helix regions were scattered in the GAP sequences of lamprey GnRH-I. The GAP sequence of different families of GnRH is assumed to have completely changed during the evolution of GnRH precursors. As the GAP sequence is conserved within the same families of GnRH precursors, the rearrangement of GAP sequence probably happened when gene duplication occurred to create a gene encoding a novel form of GnRH.

The signal peptide of lamprey GnRH-I shared an average of 36% identity with that of other GnRHs. Significantly greater identity (56%) was found

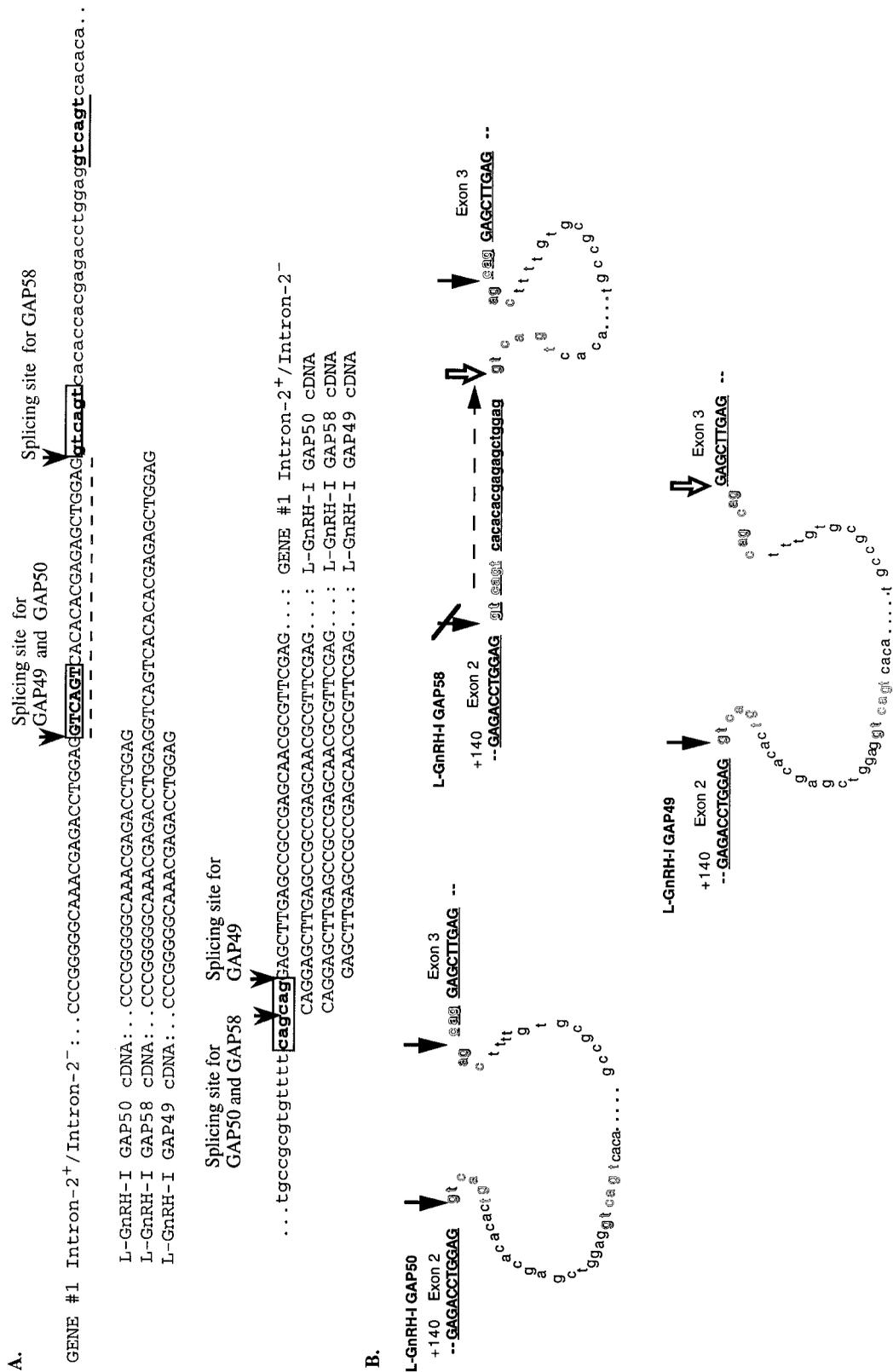


FIGURE 5. Sequence at the junction of splicing sites of lamprey GnRH-I genomic DNA. (A) Sequences of genomic DNA at the junction of the splice sites for intron-2 are aligned with that of three different lamprey GnRH-I transcripts. Possible splice signals, GTCAGT, at the beginning of intron-2 are enclosed in squares. Tandem splice signals, CAG/CAG, at the end of intron-2 are also enclosed in squares. Splice sites for each transcript are indicated by vertical arrows. The additional splice signal located at the 3'-splice junction is underlined. (B) Diagram showing the splicing of intron-2 for three different lamprey GnRH-I transcripts. Solid vertical arrows represent the splice site for the dominant transcript, lamprey GnRH-I GAP50. The solid vertical arrow with slashed line indicates the splice sequence for GAP50 that is not used in GAP58. Open arrows indicate the alternative splicing site. The exon sequences are underlined. The hollow letters indicate splice signals.

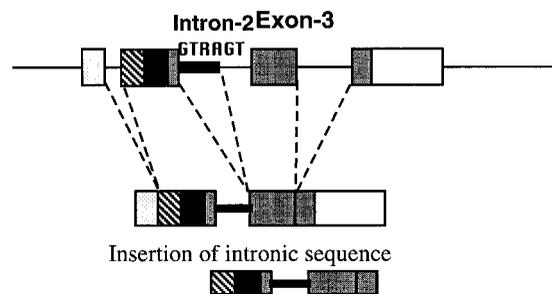
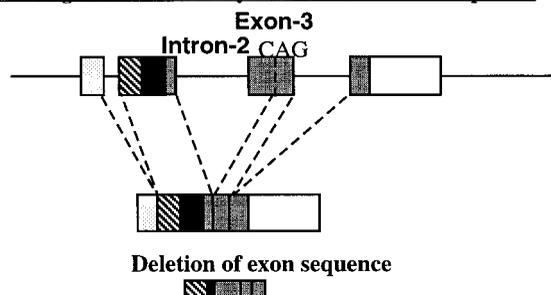
Rearrangement of GAP by insertion of Intron-2 sequence**Rearrangement of GAP by deletion of Exon-3 sequence**

FIGURE 6. Diagrams showing the possible mechanisms of rearrangement of the GAP sequence. The splice junctions of intron-2 may contribute to remodeling of the GAP structure.

between the signal peptides of lamprey GnRH-I and chicken GnRH-II from human. The first 24 residues make up the signal peptide, which is similar to known GnRH signal peptides in length and degree of hydrophobicity. However, the signal peptide of lamprey GnRH-I has no significant sequence homology (an average of 20% identity) with those of tetrapod and teleost GnRHs. Furthermore, the recognition site for signal peptidase cleavage does not appear to be conserved among the different forms of GnRH cDNA.

Analysis of lamprey genomic DNA revealed that lamprey GnRH-I is encoded by a single gene suggesting that the three GnRH-I transcripts are splice variants of mRNA. Lamprey GnRH-I is the first GnRH gene reported to contain multiple splice sequences that produce splice variants. Another study did show two nucleotide sequences for one GnRH form in the catfish (Bogerd *et al.* 1994). However, analysis of genomic DNA was not performed in these studies, and because of the small number of changes in the sequence the authors concluded that the different nucleotide sequences indicated population polymorphism rather than suggesting the presence of two catfish GnRH genes. In the present study, the differences between lamprey GnRH-I GAP49, GAP50 and GAP58 are believed to be due to alternate splicing at the intron-2 junction of the lamprey GnRH-I gene.

L-GnRH-I GAP49	MALRGQSLTLLLLLATALLVSLNYA	QHYSLEWKPG	GKRDLE	-----	EL
L-GnRH-I GAP50	MALRGQSLTLLLLLATALLVSLNYA	QHYSLEWKPG	GKRDLE	-----	OEL
L-GnRH-I GAP58	MALRGQSLTLLLLLATALLVSLNYA	QHYSLEWKPG	GKRDLE	--VSH	TRELEQEL
chicken GnRH-II	MCVSRLALLLGLLLCVG-AQLSFA	QHW SHGWYPG	GKRELD	SFGT	---EISEEI
seabream GnRH	MAAKILALWL-LLAGIV-FPQCC	QHW SYGLSPG	GKRDLD	NFSD	ILGNMVEEF
salmon GnRH	MEAGSRVIMOVLLALV-VQVTL	QHW SYGWLPG	GKRSVGE	LEAT	IRMMGT--
chicken GnRH-I	MEKSRKILVGVLLFTAS-VATCLA	QHW SYGLQPG	GKRNAEN	LVES	FQETANEM
mammalian GnRH	MKPIQKLLAGLILLTSC-VEGCCS	QHW SYGLRPG	GKRDAEN	LID	SFQETVKEV
L-GnRH-I GAP49	EPPSNAFECDGPECA--FSRVPNT	KLIRELAS	YLSQR	----	NYDRKGAL--K
L-GnRH-I GAP50	EPPSNAFECDGPECA--FSRVPNT	KLIRELAS	YLSQR	----	NYDRKGAL--K
L-GnRH-I GAP58	EPPSNAFECDGPECA--FSRVPNT	KLIRELAS	YLSQR	----	NYDRKGAL--K
chicken GnRH-II	KL-----	CEAGECSYLRP-QRRSILRN	ILLDALAREL	---	Q-KRK
seabream GnRH	PRVEAP--	CSVFGCAEESPF	AKMYRVKGLLASVAERKMDTGH-SRNERFL		
salmon GnRH	-----	GGVVS	LPDEANAQIQERLR	PYNI	INDDSSHFD
chicken GnRH-I	ESLGEG---	QKA	ECPGSYQHPRLSDLKETMASL	----	TEGEA-RRKE--I
mammalian GnRH	GQLAET---	QR	FECTTHQPRSPRLDLKGALES	----	IEEET-GQKK--I

FIGURE 7. Comparison of GnRH precursor sequences. The sequences of prepro-GnRH were aligned by MegAlign (DNA STAR) using the PAM250 weight table. Boxed residues indicate the GnRH decapeptide sequence. Chicken GnRH-II, seabream GnRH and salmon GnRH sequences are all from cichlid (Bond *et al.* 1991, White *et al.* 1994, 1995). The chicken GnRH-I sequence is from chicken (Dunn *et al.* 1993), and the mammalian GnRH is from human (Adelman *et al.* 1986).

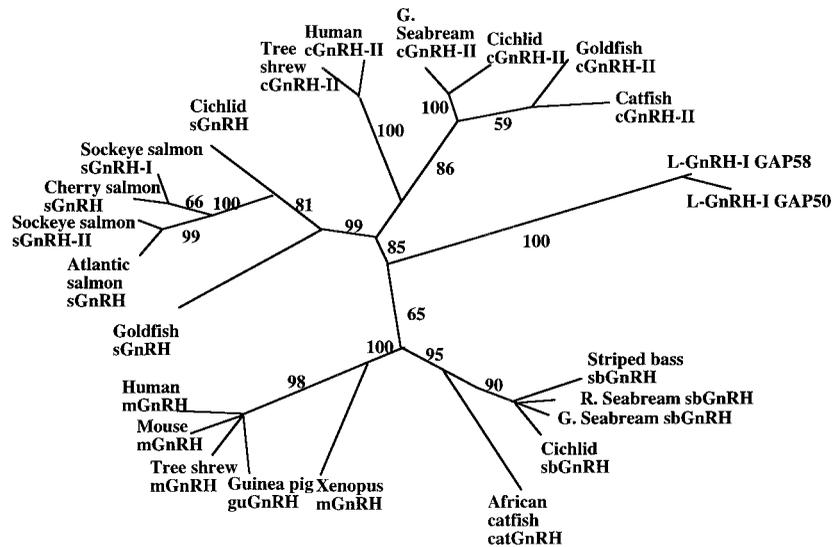


FIGURE 8. Phylogenetic tree generated by the cDNA coding prepro-GnRHs. The sequences of prepro-GnRH were aligned by MegAlign. The tree was generated by distance matrix method using PAUP 4.0 (Swofford 1998). Bootstrap values (confidence values) are shown at the nodes on the tree and indicate the number of times the same branching point is selected in 100 replicates. sGnRH, Salmon GnRH; cGnRH-I, chicken GnRH-I; cGnRH-II, chicken GnRH-II; sbGnRH, seabream GnRH; catGnRH, catfish GnRH; mGnRH, mammalian GnRH; R. Seabream, red seabream; G. Seabream, gilthead seabream.

Both the 5'-splice recognition sequence GTCAGT and the 3'-splice recognition sequence CAG are found in duplicate near the intron-2 splice junction. The splice sites of GnRH genes used in gnathostomes and in lamprey GnRH-I GAP58 are located downstream of the splice site of lamprey GnRH-I GAP50.

The 5'-splice sequence of the lamprey GnRH-I gene resembles the mammalian consensus sequence of GT(A/G)AGT, except that the lamprey sequence substitutes a C in the third position. Also similar to the mammalian GnRH gene, the 3'-splice junction of intron-2 of lamprey GnRH-I gene has a typical pyrimidine tract that is rich in the nucleotides T and C. The consensus sequence CAG was found downstream of such a pyrimidine tract in the lamprey GnRH-I. In contrast to the variable splice sequences of intron-2, the splice junctions of intron-3 are conserved among the lamprey GnRH-I transcripts. The 5'-splice sequences of intron-3 in all known GnRH genes, including lamprey, are highly homologous to the general consensus splice sequence, GT(A/G)GT. In addition, all 3'-splice sequences of intron-3 in known GnRH genes are homologous to the consensus (T/C)AG in mammalian genes.

In the present study, the tissue distribution of lamprey GnRH-I mRNA showed primary expression of lamprey GnRH-I in the brain, with slight expression in the testis and none in any other tissues examined. These data support immunocytochemical and physiological studies demonstrating lamprey GnRH-I immunoreactivity in cell bodies of the rostral hypothalamus and the preoptic area in larval and adult sea lamprey in addition to demonstrating GnRH action on the pituitary-gonadal axis (Sower 1998). GnRH expression in non-neural tissues in other organisms is documented for the uterus, ovary, testis, breast, pituitary, spleen, kidney, thymus, and blood lymphocyte (Azad *et al.* 1991, Palmon *et al.* 1994, Kakar & Jennes 1995, Dong *et al.* 1998, White & Fernald 1998). The function of GnRH in non-neural and non-gonadal tissues is unknown. However, it has been speculated that GnRH acts as a local paracrine regulatory factor in gonads of teleost fish (Habibi *et al.* 1988, 1989). Previous studies determined that lamprey GnRH could act directly on gonads to stimulate steroidogenesis (Gazourian *et al.* 1997) and on the basis of these data, GnRH is presumed to act as paracrine modulator of lamprey gonads.

Phylogenetic analysis of the precursors (both the amino acid sequences of prepro-GnRHs and the cDNA coding prepro-GnRHs) showed the same GnRH precursors grouped together in distinct clades, which can be seen for the salmon GnRH precursors, chicken GnRH-II precursors, seabream GnRH precursors and mammalian GnRH precursors. Two exceptions include the chicken GnRH-II precursors isolated from the tree shrew and from the human that do not appear to be similar to the other chicken GnRH-II precursors. This is due to the absence of a large section of sequence encoding for GAP. The second exception is the mammalian GnRH precursor from the frog, *Xenopus laevis*. This sequence appears to be as different from the other mammalian GnRH precursors as is the chicken GnRH-I precursor sequence from mammalian GnRH. It can be postulated that with the identification of more chicken GnRH-I precursors from other animals, the chicken GnRH-I clade will form its own branch and this will no longer be the case. Because of the evolutionary distance between lampreys and gnathostomes, the position of lamprey GnRH-I precursor is distantly located from all other GnRH branches. As a significant degree of similarity was conserved in the same family of GnRH precursors, each GnRH family formed their own branch.

A recent phylogenetic study by White *et al.* (1998) showed that GnRH could be classified into three families: mesencephalic forms, hypothalamic-releasing forms and telecephalic forms. According to the findings of that study, seabream GnRH belongs to the mammalian GnRH family. In contrast, phylogenetic analyses in the present study showed each family having its own branch. To date, the precursor sequences of the more primitive GnRH precursors – dogfish and tunicate-I and -II – and lamprey GnRH-III are not yet known. The precursor sequences from these forms, and those from novel GnRH forms not yet identified, will aid in further resolving the phylogenetic relationships of the GnRH family. Characterization of brain hormone genes from an extant representative of the oldest class of vertebrates is particularly important for understanding the molecular evolution and functional diversity of these hormones.

ACKNOWLEDGEMENTS

The authors would like to thank Julie Heinig and John Youson for their guidance early in this project. Thanks also go to Robin Rosenbleeth and Salvatore Scali for laboratory assistance. We especially thank Drs Karen L Reed and Jill B K Leonard for their

assistance and editing of this manuscript. This research was funded by NSF-MCB 9722765, NSF-FAW-9022834, the Great Lakes Fishery Commission and the UNH Hubbard Marine Program Endowment. Scientific contribution No. 2025 from the New Hampshire Agricultural Experiment Station.

REFERENCES

- Adelman JP, Mason AJ, Hayflick JS & Seeburg PH 1986 Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat. *Proceedings of the National Academy of Sciences of the USA* **83** 179–183.
- Andersen O & Klungland H 1993 The salmon GnRH encoding gene in teleost fish. *International Reviews in Cytology* **147** 165–191.
- Ashihara M, Suzuki M, Kubokawa K, Yoshiura Y, Kobayashi M, Urano A & Aida K 1995 Two differing precursor genes for the salmon-type gonadotropin-releasing hormone exist in salmonids. *Journal of Molecular Endocrinology* **15** 1–9.
- Azad N, Emanuele NV, Halloran MM, Tentler J & Kelley MR 1991 Presence of luteinizing hormone-releasing hormone (LHRH) mRNA in rat spleen lymphocytes. *Endocrinology* **128** 1679–1681.
- Bogerd J, Zandbergen T, Andersson E & Goos H 1994 Isolation, characterization and expression of cDNAs encoding the catfish-type and chicken-II-type gonadotropin-releasing-hormone precursors in the African catfish. *European Journal of Biochemistry* **222** 541–549.
- Bond CT, Francis RC, Fernald RD & Adelman JP 1991 Characterization of complementary DNA encoding the precursor for gonadotropin-releasing hormone and its associated peptide from a teleost fish. *Molecular Endocrinology* **5** 931–937.
- Chavali GB, Nagpal S, Majumdar SS, Singh O & Salunke DM 1997 Helix-loop-helix motif in GnRH associated peptide is critical for negative regulation of prolactin secretion. *Journal of Molecular Biology* **272** 731–740.
- Chomzynski P & Sacchi N 1987 Single step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analytical Biochemistry* **162** 156–159.
- Chow MM, Kight KE, Gothilf Y, Alok D, Stubblefield J & Zohar Y 1998 Multiple GnRHs present in a teleost species are encoded by separate genes: analysis of the sbGnRH and cGnRH-II genes from the striped bass, *Morone saxatilis*. *Journal of Molecular Endocrinology* **21** 277–289.
- Coe IR, von Schalburg KR & Sherwood NM 1995 Characterization of the Pacific salmon gonadotropin-releasing hormone gene, copy number and transcription start site. *Molecular and Cellular Endocrinology* **115** 113–122.
- Dong KW, Yu KL & Roberts JL 1993 Identification of a major up-stream transcription start site for the human progonadotropin-releasing hormone gene used in reproductive tissues and cell lines. *Molecular Endocrinology* **7** 1654–1666.
- Dong KW, Marcelin K, Hsu MI, Chiang CM, Hoffman G & Roberts JL 1998 Expression of gonadotropin-releasing hormone (GnRH) gene in human uterine endometrial tissue. *Molecular and Human Reproduction* **4** 893–898.
- Dunn IC, Chen Y, Hook C, Sharp PJ & Sang HM 1993 Characterization of the chicken

- preprogonadotrophin-releasing hormone-I gene. *Journal of Molecular Endocrinology* **11** 19–29.
- Frohman MA, Dush MK & Martin GR 1988 Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences of the USA* **85** 8998–9002.
- Garnier J & Robson B 1989 *Prediction Of Protein Structure And Principles Of Protein Conformation*. New York: Plenum.
- Gazourian L, Deragon KL, Chase CF, Pati D, Habibi HR & Sower SA 1997 Characteristics of GnRH binding in the gonads and effects of lamprey GnRH-I and -III on reproduction in the adult sea lamprey. *General and Comparative Endocrinology* **108** 327–339.
- Gothilf Y, Elizur A, Chow M, Chen TT & Zohar Y 1995 Molecular cloning and characterization of a novel gonadotropin-releasing hormone from the gilthead seabream (*Sparus aurata*). *Molecular Marine Biology and Biotechnology* **4** 27–35.
- Gothilf Y, Munoz-Cueto JA, Sagrillo CA, Selmanoff M, Chen TT, Kah O, Elizur A & Zohar Y 1996 Three forms of gonadotropin-releasing hormone in a perciform fish (*Sparus aurata*): complementary deoxyribonucleic acid characterization and brain localization. *Biology of Reproduction* **55** 636–645.
- Goubau S, Bond CT, Adelman JP, Misra V, Hynes MF, Schultz GA & Murphy BD 1992 Partial characterization of the gonadotropin-releasing hormone (GnRH) gene transcript in the rat ovary. *Endocrinology* **130** 3098–3100.
- Grober MS, Myers TR, Marchaterre MA, Bass AH & Myers DA 1995 Structure, localization, and molecular phylogeny of a GnRH cDNA from a paracanthopterygian fish, the plainfin midshipman (*Porichthys notatus*). *General and Comparative Endocrinology* **99** 85–99.
- Habibi HR, Van der Kraak G, Bulanski E & Peter RE 1988 Effect of teleost GnRH on reinitiation of oocyte meiosis in goldfish, *in vitro*. *American Journal of Physiology* **255** R268–R273.
- Habibi HR, Van Der Kraak G, Fraser R & Peter RE 1989 Effect of a teleost GnRH analog on steroidogenesis by the follicle-enclosed goldfish oocytes, *in vitro*. *General and Comparative Endocrinology* **76** 95–105.
- Hayes WP, Wray S & Battey JF 1994 The frog gonadotropin-releasing hormone-I (GnRH-I) gene has a mammalian-like expression pattern and conserved domains in GnRH-associated peptide, but brain onset is delayed until metamorphosis. *Endocrinology* **134** 1835–1845.
- Hayflick JS, Adelman JP & Seeburg PH 1989 The complete nucleotide sequence of the human gonadotropin-releasing hormone gene. *Nucleic Acids Research* **17** 6403–6404.
- Higa M, Kitahashi T, Sasaki Y, Okada H & Ando H 1997 Distinct promoter sequences of two precursor genes for salmon gonadotropin-releasing hormone in masu salmon. *Journal of Molecular Endocrinology* **19** 149–161.
- Jimenez-Linan M, Rubin BS & King JC 1997 Examination of guinea pig luteinizing hormone-releasing hormone gene reveals a unique decapeptide and existence of two transcripts in the brain. *Endocrinology* **138** 4123–4130.
- Kakar SS & Jennes L 1995 Expression of gonadotropin-releasing hormone and gonadotropin-releasing hormone receptor mRNAs in various non-reproductive human tissues. *Cancer Letters* **98** 57–62.
- Klungland H, Andersen O, Kisen G, Alestrom P & Tora L 1992a The salmon gonadotropin-releasing hormone encoding gene in salmonids. *Molecular Marine Biology and Biotechnology* **1** 420–425.
- Klungland H, Lorens JB, Andersen O, Kisen GO & Alestrom P 1992b The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA. *Molecular and Cellular Endocrinology* **84** 167–174.
- Kyte J & Doolittle RF 1982 A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology* **157** 105–132.
- Leader DP, Gall I, Campbell P & Frischauf AM 1986 Isolation and characterization of cDNA clones from mouse skeletal muscle actin mRNA. *DNA* **5** 235–238.
- Lin XW & Peter RE 1996 Expression of salmon gonadotropin-releasing hormone (GnRH) and chicken GnRH-II precursor messenger ribonucleic acids in the brain and ovary of goldfish. *General and Comparative Endocrinology* **101** 282–296.
- Mason AJ, Hayflick JS, Zoeller RT, Young WSD, Phillips HS, Nikolics K & Seeburg PH 1986 A deletion truncating the gonadotropin-releasing hormone gene is responsible for hypogonadism in the hpg mouse. *Science* **234** 1366–1371.
- Nielsen H, Engelbrecht J, Brunak S & von Heijne G 1997 Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* **10** 1–6.
- Nikolics K, Mason AJ, Szonyi E, Ramachandran J & Seeburg PH 1985 A prolactin inhibiting-factor within the precursor for human gonadotropin-releasing hormone. *Nature* **316** 511–517.
- Okuzawa K, Araki K, Tanaka H, Kagawa H & Hirose K 1994 Molecular cloning of a cDNA encoding the prepro-salmon gonadotropin-releasing hormone of the red seabream. *General and Comparative Endocrinology* **96** 234–242.
- Palmon A, Ben Aroya N, Tel-Or S, Burstein Y, Fridkin M & Koch Y 1994 The gene for the neuropeptide gonadotropin-releasing hormone is expressed in the mammary gland of lactating rats. *Proceedings of the National Academy of Sciences of the USA* **91** 4994–4996.
- Planas J, Bern HA & Millar RP 1990 Effects of GnRH-associated peptide and its component peptides on prolactin secretion from the tilapia pituitary *in vitro*. *General and Comparative Endocrinology* **77** 386–396.
- Radovick S, Wondisford FE, Nakayama Y, Yamada M, Cutler GB Jr & Weintraub BD 1990 Isolation and characterization of the human gonadotropin-releasing hormone gene in the hypothalamus and placenta. *Molecular Endocrinology* **4** 476–480.
- Sambrook J, Fritsch EF & Maniatis RP 1989 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.
- Seeburg PH & Adelman JP 1984 Characterization of cDNA for precursor of human luteinizing hormone releasing hormone. *Nature* **311** 666–668.
- Sherwood NM, Sower SA, Marshak DR, Fraser BA & Brownstein MJ 1986 Primary structure of gonadotropin-releasing hormone from lamprey brain. *Journal of Biological Chemistry* **261** 4812–4819.
- Sonnhammer ELL, von Heijne G & Krogh A 1998 A hidden Markov model for predicting transmembrane helices in protein sequences. *ISMB* **6** 175–182.
- Sower SA 1998 Brain and pituitary hormones of lampreys, recent findings and their evolutionary significance. *American Zoologist* **38** 15–38.
- Sower SA, Chiang YC, Lovas S & Conlon JM 1993 Primary structure and biological activity of a third gonadotropin-releasing hormone from lamprey brain. *Endocrinology* **132** 1125–1131.
- Suzuki M, Hyodo S, Kobayashi M, Aida K & Urano A 1992 Characterization and localization of mRNA encoding the salmon-type gonadotrophin-releasing hormone precursor of the masu salmon. *Journal of Molecular Endocrinology* **9** 73–82.
- Swofford DL 1998 *PAUP: Phylogenetic Analysis Using Parsimony*. Champaign, IL: Illinois Natural History Survey.

- Tautz D & Renz M 1983 An optimized freeze-squeeze method for the recovery of DNA fragments from agarose gels. *Analytical Biochemistry* **132** 14–19.
- Thomas GB, Cummins JT, Doughton BW, Griffin N, Millar RP, Milton RC & Clarke IJ 1988 Gonadotropin-releasing hormone associated peptide (GAP) and putative processed GAP peptides do not release luteinizing hormone or follicle-stimulating hormone or inhibit prolactin secretion in the sheep. *Neuroendocrinology* **48** 342–350.
- Van Chuoit MT, Vacher P & Dufy B 1993 GnRH-associated peptide decreases cyclic AMP accumulation in the GH3 pituitary cell line. *Neuroendocrinology* **58** 251–257.
- White RB & Fernald RD 1998 Genomic structure and expression sites of three gonadotropin-releasing hormone genes in one species. *General and Comparative Endocrinology* **112** 17–25.
- White SA, Bond CT, Francis RC, Kasten TL, Fernald RD & Adelman JP 1994 A second gene for gonadotropin-releasing hormone: cDNA and expression pattern in the brain. *Proceedings of the National Academy of Sciences of the USA* **91** 1423–1427.
- White SA, Kasten TL, Bond CT, Adelman JP & Fernald RD 1995 Three gonadotropin-releasing hormone genes in one organism suggest novel roles for an ancient peptide. *Proceedings of the National Academy of Sciences of the USA* **92** 8363–8367.
- White RB, Eisen JA, Kasten TL & Fernald RD 1998 Second gene for gonadotropin-releasing hormone in humans. *Proceedings of the National Academy of Sciences of the USA* **95** 305–309.

RECEIVED 28 September 1999