0013-7227/03/$15.00/0
Printed in U.S.A.

Endocrinology 144(2):423–436

doi: 10.1210/en.2002-220622

A Transcriptionally Active Human Type II Gonadotropin-Releasing Hormone Receptor Gene Homolog Overlaps Two Genes in the Antisense Orientation on Chromosome 1q.12

KEVIN MORGAN, DARRELL CONKLIN, ADAM J. PAWSON, ROBIN SELLAR, THOMAS R. OTT, AND ROBERT P. MILLAR

Medical Research Council Human Reproductive Sciences Unit (K.M., A.J.P., R.S., T.R.O., R.P.M.), University of Edinburgh Academic Centre, Edinburgh EH16 4SB, United Kingdom; ZymoGenetics Inc. (D.C.), Seattle, Washington 98102; and Medical Research Council Molecular Reproductive Endocrinology Unit (R.P.M.), University of Cape Town Medical School, Observatory 7925, Cape Town, South Africa

GnRH-II peptide hormone exhibits complete sequence conservation across vertebrate species, including man. Type-II GnRH receptor genes have been characterized recently in nonhuman primates, but the human receptor gene homolog contains a frameshift, a premature stop codon (UGA), and a 3′ overlap of the RBM8A gene on chromosome 1q.12. A retrotransposed pseudogene, RBM8B, retains partial receptor sequence. In this study, bioinformatics show that the human receptor gene promoter overlaps the peroxisomal protein11-β gene promoter and the premature UGA is positionally conserved in chimpanzee. A CGA [arginine (Arg)] occurs in porcine DNA, but UGA is shifted one codon to the 5′ direction in bovine DNA, suggesting independent evolution of premature stop codons. In contrast to marmoset tissue RNA, exon- and strand-specific probes are required to distinguish differently spliced human receptor gene transcripts in cell lines (HP75, IMR-32). RBM8B is not transcribed. Sequencing of cDNAs for spliced receptor mRNAs showed no evidence for alteration of the premature UGA by RNA editing, but alternative splicing circumvents the frameshift to encode a two-membrane-domain protein before this UGA. A stem-loop motif resembling a selenocysteine insertion sequence and a potential alternative translation initiation site might enable expression of further proteins involved in interactions within the GnRH system. (*Endocrinology* 144: 423–436, 2003)

THE SECRETION of gonadotropins, LH, and FSH is stimulated by GnRH, by binding to a seven-transmembrane G protein-coupled receptor (GPCR) known as the type I GnRH receptor. This decapeptide hormone exhibits structural variation across vertebrate species (1, 2). A homologous (yet distinct) GnRH, known as type II GnRH, is present in the majority of vertebrates, and this hormone shows complete structural conservation from teleosts to mammals (3, 4). This conservation suggests that it serves an important physiological role (5) through interaction with GnRH receptors in different tissues.

The physiology and histology of GnRH precursor peptide and GnRH receptor distribution have been studied in pituitary, brain, and reproductive tissues (6–11); and the effects of GnRH analogs on reproductive function and animal behavior have been translated into therapeutic applications (12–14).

Molecular details concerning the regulation of GnRH and GnRH receptor genes continue to emerge. Such areas of study have involved the characterization of gene regulation in murine gonadotrope and human neuronal or reproductive tissue cell lines (15–18). These studies should enable a deeper understanding of the clinical features of human infertility, certain inherited conditions (19), and of neoplasias (20–22) affecting the human reproductive system.

Gene cloning revealed a highly selective receptor for type II GnRH in amphibia (23) and subsequently in the marmoset monkey (*Callithrix jacchus*) (5), rhesus monkey (*Macaca mulatta*), and African green monkey (*Cercopithecus aethiops*) (24). The primate type II GnRH receptor shows 68% amino acid sequence similarity (40% identity) to the type I GnRH receptor; but, unlike the type I receptor, the type II GnRH receptor has a carboxyterminal cytoplasmic tail domain potentially able to influence the intracellular-recycling properties of the protein (5).

More recently, human type II GnRH receptor-like cDNAs and genomic sequences on chromosome 1q.12 that overlap those of another gene, RBM8A, have been described but not fully characterized (GenBank accession no. AF403014 and accession no. AY081843) (25–29). This human receptor gene homolog displays important differences when compared with the other primate sequences, in that the protein coding sequences contain a frameshift and a premature stop codon.

Here we extend the analysis of the genomic sequence encoding the human type II GnRH receptor gene homolog and its flanking sequences. We show that the type II GnRH receptor gene is closely linked to the gene encoding perox-

Abbreviations: Arg, Arginine; BLAST, Basic Local Alignment Search Tool; EST, expressed sequence tag; HTG, high throughput genomic; GPCR, G-protein-coupled receptor; MOPS, 4-morpholinopropanesulfonic acid; ORF, open reading frame; SDS, sodium dodecyl sulfate; SECIS, selenocysteine incorporation sequence; SSPE, standard saline phosphate-EDTA; TAE, Tris-acetate-EDTA; TMD, transmembrane domain; UTR, untranslated region.

isomal membrane protein 11-β (PEX11β) (30). The distal promoter region of the receptor gene overlaps that of the PEX11β gene, and the 3′ untranslated region (UTR) of RBM8A gene transcripts overlap with the two C-terminal coding exons of the type II receptor gene (25–29).

Using RT-PCR assays and Northern blotting, we demonstrate that a variety of alternatively spliced low-abundance transcripts are derived from the type II GnRH receptor gene in different human cell types, similar to the situation observed with GHRH receptor (31, 32), a bullfrog GnRH receptor subtype (33), and the human type I GnRH receptor (34).

Significantly, alternative splicing of human type II GnRH receptor transcripts bypasses the frameshift, while retaining the premature stop codon. In view of its potential importance in reproductive physiology, we clarify and extend the analysis of mechanisms through which functional proteins might be generated from this gene (27). In addition, we present a study of partial sequences encoding the chimpanzee, porcine, and bovine type II GnRH receptors in the region containing the premature stop codon and discuss its significance and likely evolutionary origins.

Materials and Methods

Bioinformatics studies

Type II GnRH receptor cDNA sequences. A full-length type II GnRH receptor cDNA was cloned recently from the common marmoset monkey (GenBank accession no. AF368286; Ref. 5), rhesus monkey (GenBank accession no. AF353987), and African green monkey (GenBank accession no. AF353988; Ref. 24). These cDNA sequences were used to search for similar sequences derived from other species deposited in the GenBank database.

DNA database searches. The National Center for Biotechnology Information GenBank DNA databases were screened on-line using NIH Basic Local Alignment Search Tool (BLAST) software (35).

DNA contig assembly. Consensus assembly of expressed sequence tag (EST) data was performed using the Rex EST assembly program (36). Elucidation of exon structure and chromosome 1q.12, chromosome 14q.22-23, and chromosome 11 DNA sequence homology was through dot plot analysis and visual inspection using the Dotter program (37), and by on-line pairwise BLAST alignment or GeneJockey II sequence alignment software (Biosoft, Cambridge, UK).

Scanning for sequence motifs in gene promoter regions. Promoter-inspector software (38) (Genomatix GmbH) was used to identify potentially important DNA sequences involved in transcription factor binding to the 5′ flanking regions for each gene identified.

Localization of alternative translation start sites. Open reading frames (ORFs) located within DNA sequence fragments were identified using National Center for Biotechnology Information ORF-Finder software according to information regarding translation initiation sites in eukaryotes (39). Hydropathy profiles of protein secondary structure were analyzed using PHDhtm-prediction software (EMBnet) (40).

RNA secondary structure analysis. Models of potential stem-loop structures within short stretches (50–70 bp) of gene transcripts were constructed using M-fold version 3.1 software (Rensselaer Institute, Troy, NY) (41, 42).

Phylogenetic analysis. Genomic DNA fragments from different animal species with homology to the marmoset type II GnRH receptor were translated into peptide sequences and were aligned using GeneJockey II software. Phylogenetic trees were constructed from the aligned peptides using PHYLIP software (EMBnet) (43).

Gene expression studies

Tissue samples. Marmoset tissues (adrenal, liver, kidney, ovary, pituitary, testes, skeletal muscle, and brain cerebellum, cortex, putamen, occipital pole, midbrain, medulla, pons) were collected post mortem, frozen on dry ice, and stored at −80 C before purification of RNA. Human blood (40 ml) was collected from a single donor, for preparation of genomic DNA, using a kit (Nucleon; Amersham Pharmacia Biotech, Amersham, Buckinghamshire, UK).

Cell lines. The human cell lines HP-75 (pituitary adenoma), IMR-32 (neuroblastoma), JEG-3 (choriocarcinoma), KELLY (neuroblastoma), LNCaP (metastatic prostate carcinoma), MCF-7 (breast adenocarcinoma), Ishikawa (endometrial carcinoma), and Jurkat (T-cell leukemia) were obtained from the American Type Culture Collection and grown as recommended (LGC, London, UK).

Purification of total RNA and polyA+ RNA. Tissue or cultured cells were homogenized in Tri-Reagent (Sigma, Poole, UK), mixed with chloroform and centrifuged to separate aqueous and organic phases according to the method of Chomczynski and Sacchi (44). Total RNA was precipitated from the aqueous phase, using isopropanol, and centrifuged, and the pellet was washed with 70% ethanol before dissolution in 0.1% diethylpyrocarbonate (Sigma)-treated sterile water. The RNA yield was estimated using UV spectrophotometry (GeneQuant *pro* machine, Amersham Pharmacia Biotech). PolyA+ RNA was prepared from 1.25 mg total RNA using oligo-deoxythymidine-spin-column chromatography (Amersham Pharmacia Biotech).

RT-PCR analysis of gene expression. From 5 μg total RNA, cDNAs were obtained, by RT, using random primers (Promega Corp., Southampton, UK) and superscript II (Life Technologies, Inc., Paisley, UK) according to the manufacturers' instructions. PCR amplification of target cDNA sequences was optimized for each primer pair using BioTaq DNA polymerase (with KCL buffer, $MgCl_2$, and deoxynucleotide triphosphates, all from Bioline, London, UK) and an OmniGene thermocycler (Hybaid, Ashford, UK). Optimization involved titration of $MgCl_2$ concentrations from 1.0–2.0 mM, annealing temperatures from 50–60 C, and the number of amplification cycles for each primer pair.

PCR amplification products were analyzed using ethidium bromide staining of DNA separated by agarose gel electrophoresis. The identity of specific DNA bands was determined by Southern blotting DNA from agarose gels onto Hybond-N+ (Amersham Pharmacia Biotech) and probing with ^{32}-P labeled cDNA fragments (random-prime labeled using a High Prime DNA labeling Kit; Roche Diagnostics, Indianapolis, IN) or 30-mer oligonucleotide probes 3′-end labeled (using terminal transferase, Promega Corp.), as described below. Certain PCR products were also characterized by direct PCR sequencing or by subcloning into plasmid vector (TOPO cloning kit; Invitrogen, Lewes, UK) before sequencing. For subcloning, PCR products were isolated by agarose gel electrophoresis in 1× Tris-acetate-EDTA (TAE) and bands were excised from the gel and spin-eluted in Ultrafree 0.45-μm filter units (Millipore Corp., Bedford, MA) before concentration, using ethanol precipitation, in the presence of 10 μg glycogen (Roche, Lewes, UK) as a coprecipitate.

GnRH receptor oligonucleotides. The sequences of primers used for RT-PCR analysis of human type II GnRH receptor expression were: MRC862 5′TCA TCC TCC TCA GTT TCT CTC C 3′ (sense, 30 bp 5′ to start codon), MRC1001 5′ AGG AGG TCT GGG CTG GAT CA 3′ (sense, exon 1b), MRC923 5′ CTG GCT GTG GAC ATC GCA TGT 3′ (sense, exon 1c), MRC 1081 5′ CTG TTC CTG TTC CAC ACG GT 3′ (sense, exon 2a), MRC1083 5′ TCT GCT GCC ACT GAC TGC CAT 3′ (sense, exon 2b), MRC1084 5′ ACC GTG TGG AAC AGG AAC AG 3′ (antisense, exon 2a), MRC1085 5′ GAC TGG GCC AGC TCA GTG GAC C 3′ (antisense, exon 2a), MRC924 5′ ATG GCA GTC AGT GGC AGC AGA 3′ (antisense, exon 2b), and MRC869 5′ TCA GAT AGA TGT TAT AGA AAT GCC 3′ (antisense, exon 3), as indicated in Figs. 1 and 2.

The sequences of additional human type II GnRH receptor oligonucleotides used in the preparation of cDNA probes were: MRC1002 5′ TGA CTA GTA AGT CGG CGG CTG 3′ (antisense, exon 1b) and MRC781 5′ CAT TAG TTT CAG GAA CAT CAG 3′ (antisense, exon 1c).

The sequences of human type II receptor 30-mer oligonucleotides used as hybridization probes for Northern and Southern blotting ex-

periments were: MRC1003 5′ ACC ATG TCT GCA GGC AAC GGC ACC CCT TGG 3′ (exon 1a, sense) and MRC1004 (antisense of MRC1003), MRC1005 5′ GGC TGG CTG TGG ACA TCG CAT GTC GGA ACA 3′ (exon 1c, sense) and MRC1006 (antisense of MRC1005), MRC1073 5′ CTG CTG CCA CTG ACT GCC ATG GCC ATC TGC 3′ (exon 2b, sense) and MRC1074 (antisense of MRC 1073), and MRC1075 5′ CTG AAG TCC CTC CCA GCC TGA GCC ACA TCC 3′ (exon 3, sense) and MRC 1076 (antisense of MRC1075).

The sequences of primers used for RT-PCR amplification of a marmoset type II GnRH receptor fragment were: MRC925 5′ ACA TCG CAT GTC GGA CAC TCA 3′ (sense) and MRC926 5′ GCA GCA GAA AGA GGC AGC AGA 3′ (antisense).

Other oligonucleotides. Oligonucleotides for RT-PCR amplification of human housekeeping gene cDNA fragments were: KMg1 5′ CAT CAC CAT CTT CCA GGA GC 3′ (sense) and KMg2 5′ ATG CCA GTG AGC TTC CCG TT 3′ (antisense) for GAPDH, and KMtr1 5′ GAA TCC CAG CAG TTT CTT TCT 3′ (sense) and KMtr2 5′ GTC TCC ACG AGC ATA CA 3′ (antisense) for transferrin receptor.

Northern blotting. Total RNA (30 μg) or polyA+ RNA (4 μg) was size-fractionated by denaturing agarose gel (1.1%) electrophoresis in 1× 4-morpholinopropanesulfonic acid (MOPS) buffer containing 3% formaldehyde, according to standard procedures (45). Ethidium bromide was added to all samples so that the size of prominent bands or markers could be estimated after UV transillumination (RNA markers, Life Technologies, Inc.). Gels were capillary blotted onto Hybond-N+ (Amersham Pharmacia Biotech) in 6× standard saline phosphate-EDTA (SSPE) and fixed by baking at 80 C for 2 h, followed by UV cross-linking for 90 sec. Hybond-N+ filters were prehybridized by rotation in glass hybridization tubes containing 6× SSPE/5× Denhardt's solution/0.1% sodium dodecyl sulfate (SDS) for 4 h at 55 C in a Hybaid oven. All radiolabeled probes were prepared using 50 μCi α^{32}P-deoxycycidine triphosphate (300 Ci/mmol, Amersham Pharmacia Biotech) and were separated from unincorporated radionucleotide using Sephadex G50 (cDNA probes) or Sephadex G25fine (30-mer oligonucleotide probes) in microspin columns (Amersham Pharmacia Biotech). The cDNA probe templates were isolated from PCR reactions by agarose gel electrophoresis in 1× TAE, band excision from the gel, and spin-elution in Ultrafree 0.45-μm filter units (Millipore Corp.).

Purified probes were denatured by boiling for 2 min and rapid cooling on ice before addition to the prehybridized filters. The cDNA probes were hybridized to filters at 55 C, whereas 30-mer oligonucleotide probes were used at 50 C, in prehybridization buffer, with rotation overnight. Excess probe was removed by washing filters briefly in SSPE/0.1% SDS with increasing stringency (from 5× SSPE to 1× SSPE, and then two washes in 0.1× SSPE as the final step) at incubation temperatures up to 65 C. Washed filters were sandwiched in cling-film and exposed (using a Sensitize unit, Amersham Pharmacia Biotech) to preflashed XAR-5 film (Eastman Kodak Co., Hemel Hempstead, UK) at −80 C in a cassette containing two intensifying screens. After exposure, films were allowed to warm to room temperature before development in a film processor. Probes were stripped from filters by two 5-min incubations in 0.1% SDS at 95–100 C, followed by a wash in 2× SSPE before reexposure to film or long-term storage at −80 C. The sizes of specific RNA bands were estimated using a standard curve generated by measuring the migration distance of standards of known size.

Southern blotting. PCR products or restriction endonuclease-digested genomic DNA samples (enzymes and size markers from Promega Corp.) were size-separated using agarose gel electrophoresis in 1× TAE buffer. Subsequently, the double-stranded DNA was denatured by soaking the gel in fresh 0.4-M sodium hydroxide solution and then capillary blotted onto Hybond-N+. Filters were then fixed, probed, and stripped, and films were analyzed as described above. A full-length human type II GnRH receptor cDNA, incorporating 1.446 kb of the 3′UTR, was prepared for use as a probe to analyze Bgl-II-digested human genomic DNA.

Results

Bioinformatics results

Our bioinformatics analyses indicated that overlapping DNA sequences on human chromosome 1q.12 contain an apparently disrupted gene encoding a type II GnRH receptor plus the RBM8A gene and the PEX 11β gene, as illustrated in Fig. 1A and as described below. In addition, we identified partial genomic DNA sequences encoding type II GnRH receptor-like sequences for chimpanzee, porcine, and bovine species.

Characterization of the human type II GnRH receptor gene

A BLAST search of the GenBank high throughput genomic (HTG) sequence database, using marmoset type II GnRH receptor cDNA as the query sequence, identified a human chromosome 1q.12 clone encoding a putative type II GnRH receptor gene (GenBank accession no. AL160282). This human type II GnRH receptor gene is composed of at least three exons (Fig. 1A, Tables 1 and 2, and GenBank accession no. ALIGN 00036).

The type II GnRH receptor gene spans approximately 7.5 kb on chromosome 1q.12. The first coding exon is 509-bp long and encodes the first four potential transmembrane domains (TMDs I–IV) and connecting domains. The second coding exon is located 4.25 kb downstream, and it encodes TMD V and flanking domains in a stretch of 211 bp. The third coding exon is separated from the second exon by a small intron of 449 bp, and it encodes TMDs VI–VII and a cytoplasmic tail domain over a region of 419 bp (Fig. 1, A and D–F). The human cDNA sequence displays 92.5% identity to the marmoset cDNA in the region encompassing exons 1–3.

The occurrence of splice sites and the pattern of reading frame utilization is indicated in Fig. 2A. Several potential splice donor/acceptor sites and polyadenylation signal sequences occur in the 3′UTR. The nearest AATAAA (46) polyadenylation signal is positioned 1224 bp 3′ to the stop codon in exon 3, and it is located within intron 4 of the RBM8A gene (Fig. 1A).

Sequence analysis of the noncoding regions and introns of the receptor gene showed three *Alu* repeat-like fragments (similar to GenBank accession no. L35531) and a partial long interspersed element fragment (similar to LINE1.3, GenBank accession no. L19088) in intron 1 (Fig. 1A). The first *Alu* fragment in intron 1 is flanked by a $(T)_{14}G$ motif, and the third *Alu* fragment is flanked by $(T)_{14}G$ and $C(A)_{16}$ motifs. The immediate 5′ and 3′ flanking regions of the gene do not contain *Alu*- or LINE-related fragments.

As previously described (25–29), a truncated copy of the receptor gene, from which a 5′ section is missing, including the first coding exon and part of intron 1, exists on chromosome 14q.22 (GenBank accession no. AL132778) (Fig. 1B). In addition, we identified a 613-bp section of sequence on chromosome 11 (GenBank accession no. AC107240 and Table 2) matching 266-bp 5′UTR flanking sequence, part of the first coding exon (17 codons), and the whole of intron 1 of the RBM8A gene.

Southern blotting results, using a full-length human type II GnRH receptor cDNA as a probe, were consistent with the existence of one full-length type II GnRH receptor gene and

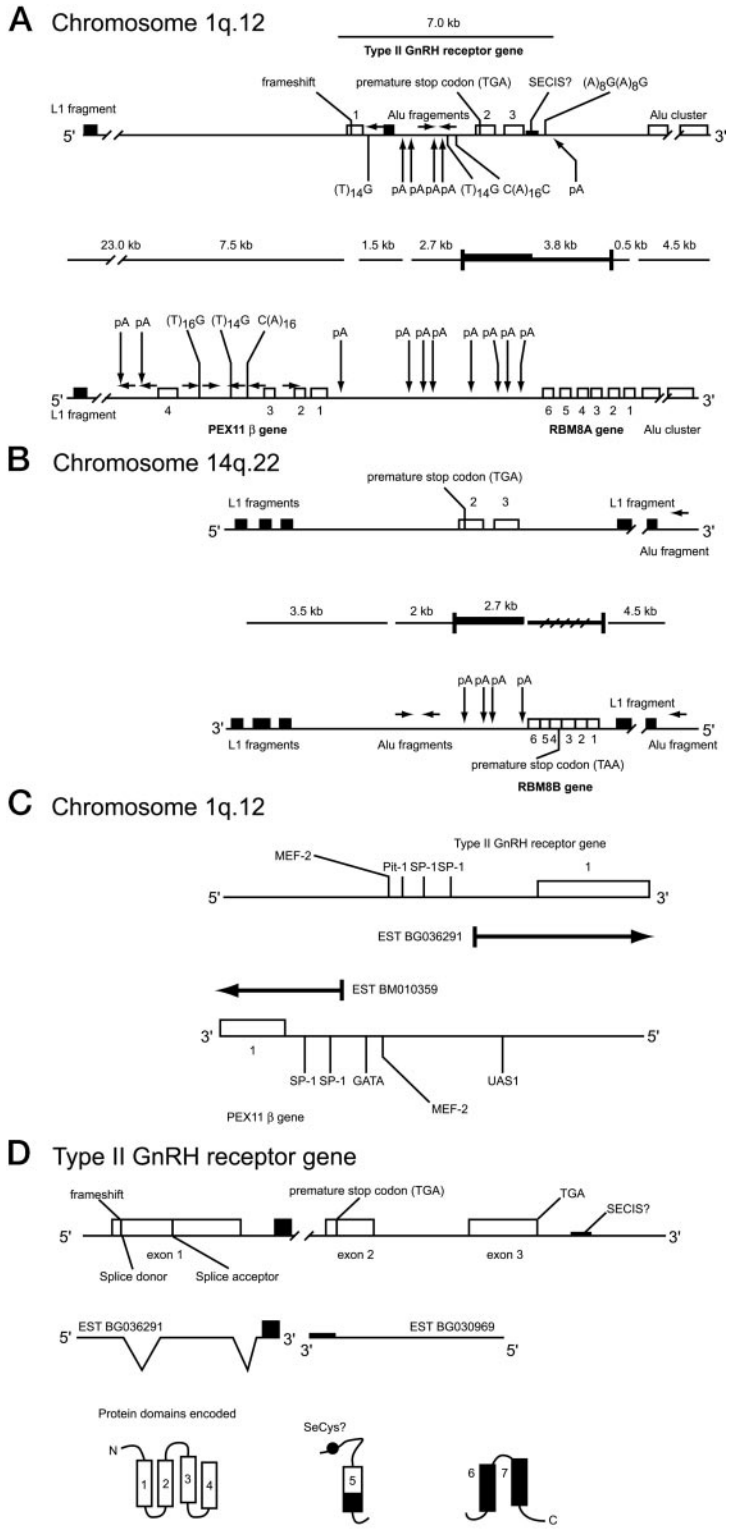

FIG. 1. A, Diagram indicating the distribution of *Alu*-fragment clusters (*open boxes not numbered*), LINE fragments (*filled boxes*), and single *Alu* fragments (*horizontal arrows*) surrounding the type II GnRH receptor (*upper line*) -RBM8A/PEX11β genes (*lower line*) on human chromosome 1q.12. Coding exons are numbered and polyadenylation signal motifs indicated (pA). The region of sequence identity (>99%) between chromosome 1q.12 and chromosome 14q.22-23 is indicated as a *thick black line* within the 3.8-kb subregion. B, Diagram of the RBM8B gene locus on human chromosome 14q.22-23; symbols as in A. C, Diagram of the region between the first coding exons of the type II GnRH receptor and PEX11β genes, indicating the relative positions of the two longest EST clones reported so far and several transcription factor consensus binding sites. D, Human type II GnRH receptor gene coding exons and their corresponding protein domains. DNA sequence alignment with EST BG036291 and EST 030969 is indicated. Note splicing into LINE fragment for EST BG036291. Partial retention of intron 1 in EST 030969 is represented in *bold*. Protein domains encoded by each exon are drawn below the DNA, where cylinders represent putative TMDs. The premature TGA stop codon is shown translated as selenocysteine (SeCys). E, Alignment of marmoset and human cDNA sequences encoding type II GnRH receptor. The positions of the frameshift, premature stop codon, and a potential alternative translation initiation site in the human sequence are indicated. The splice donor (GTCAGC) and acceptor (CCCTAG) sites within exon 1 and oligonucleotides used in this study are *underlined in bold*. F, Alignment of the predicted primary structures of the marmoset and human type II GnRH receptors. Protein domains are indicated, including those skipped by alternative splicing. The positions of the frameshift and potential internal translation initiation site are *arrowed*. The position of the premature stop codon is marked (*asterisk*). The protein encoded by EST BG036291 is indicated. G, Illustration of a putative SECIS-like stem-loop structure in the 3′UTR of the type II GnRH receptor gene (not to scale). *Arrows* indicate the four non-Watson-Crick base pairs in helix 2 and the conserved AAG/A loop.

one truncated type II GnRH receptor gene (coding exons 2 and 3) in the human genome (data not shown).

The type II GnRH receptor gene reading frame is disrupted

Comparison of the marmoset cDNA with exon 1 of the human type II receptor reveals a −1 frameshift in the human sequence located 28-bp downstream of the putative initiating ATG (Fig. 1, D–F). The frameshift is caused by the deletion of a single base (a G residue) from the human sequence relative to the marmoset sequence. Furthermore, comparison of the human type II exon 2 with the marmoset sequence reveals a C-to-T substitution that changes the codon for

FIG. 1. *Continued*

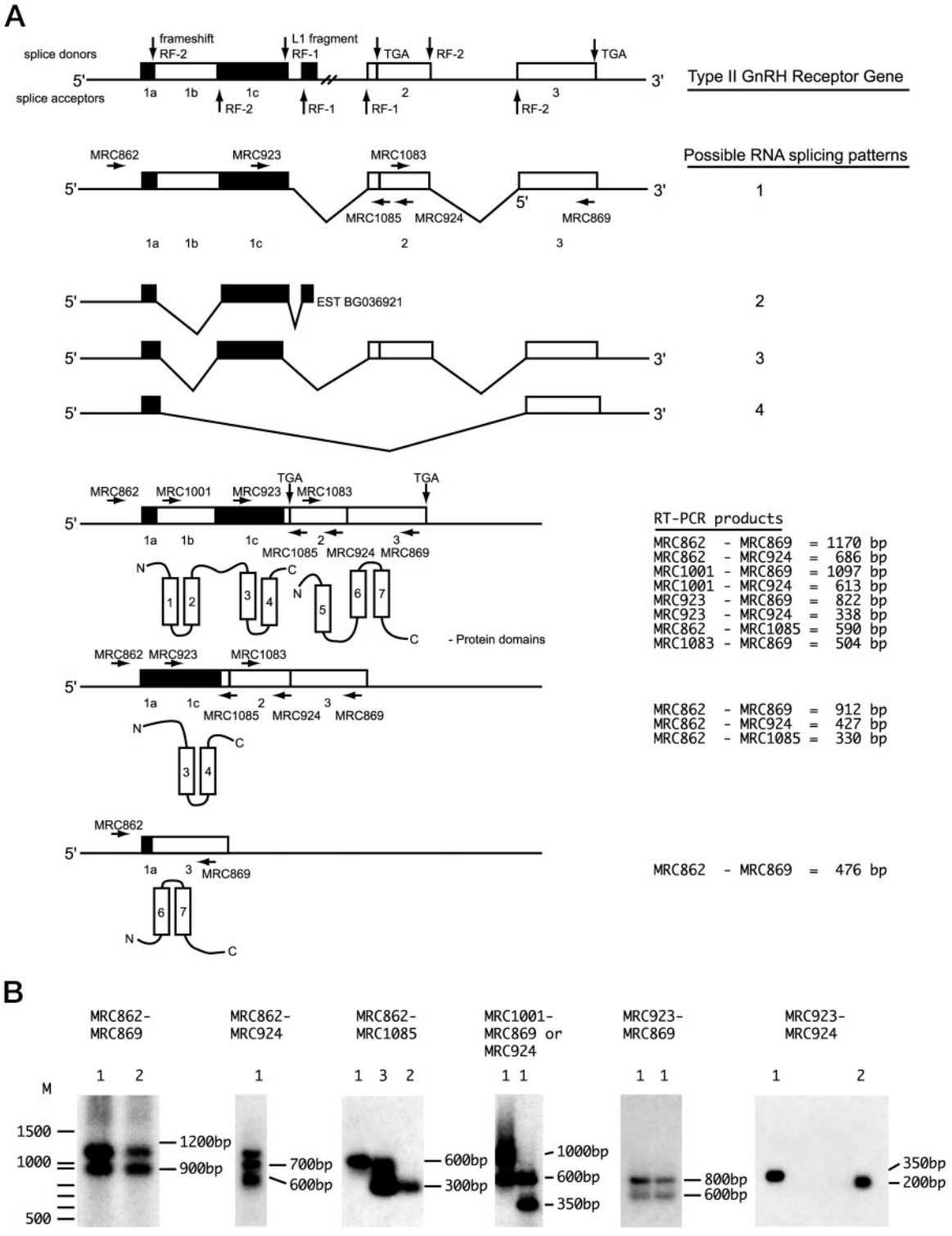

FIG. 2. A, Diagram indicating possible transcript splicing patterns for the human type II GnRH receptor gene. Exons are depicted as *boxes*. The reading frame (RF) of the sequence before each splice donor or succeeding each splice acceptor sequence is indicated (*vertical arrows*). The predicted sizes of cDNA fragments generated by combinations of PCR primers (*horizontal arrows*) are indicated alongside each primer combination. B, Examples of RT-PCR-Southern blotting on HP-75 RNA (lane 1), IMR-32 RNA (lane 2), and LNCaP RNA (lane 3). The PCR primer pairs used are indicated *above each panel*. The approximate sizes of bands hybridizing to exon 1 probe are indicated. C, Results of a polyA+ RNA Northern blot (*top panel*) compared with a Northern blot of polyA+ RNA from human cell lines both hybridized with double-stranded probes to type II GnRH receptor exon 1 and transferrin receptor. Probe-specific bands are indicated (M, RNA size markers). D, Results of further hybridizations to human mRNA Northern blot from C. The approximate sizes of transcripts detected with strand-specific oligonucleotide probes are indicated.

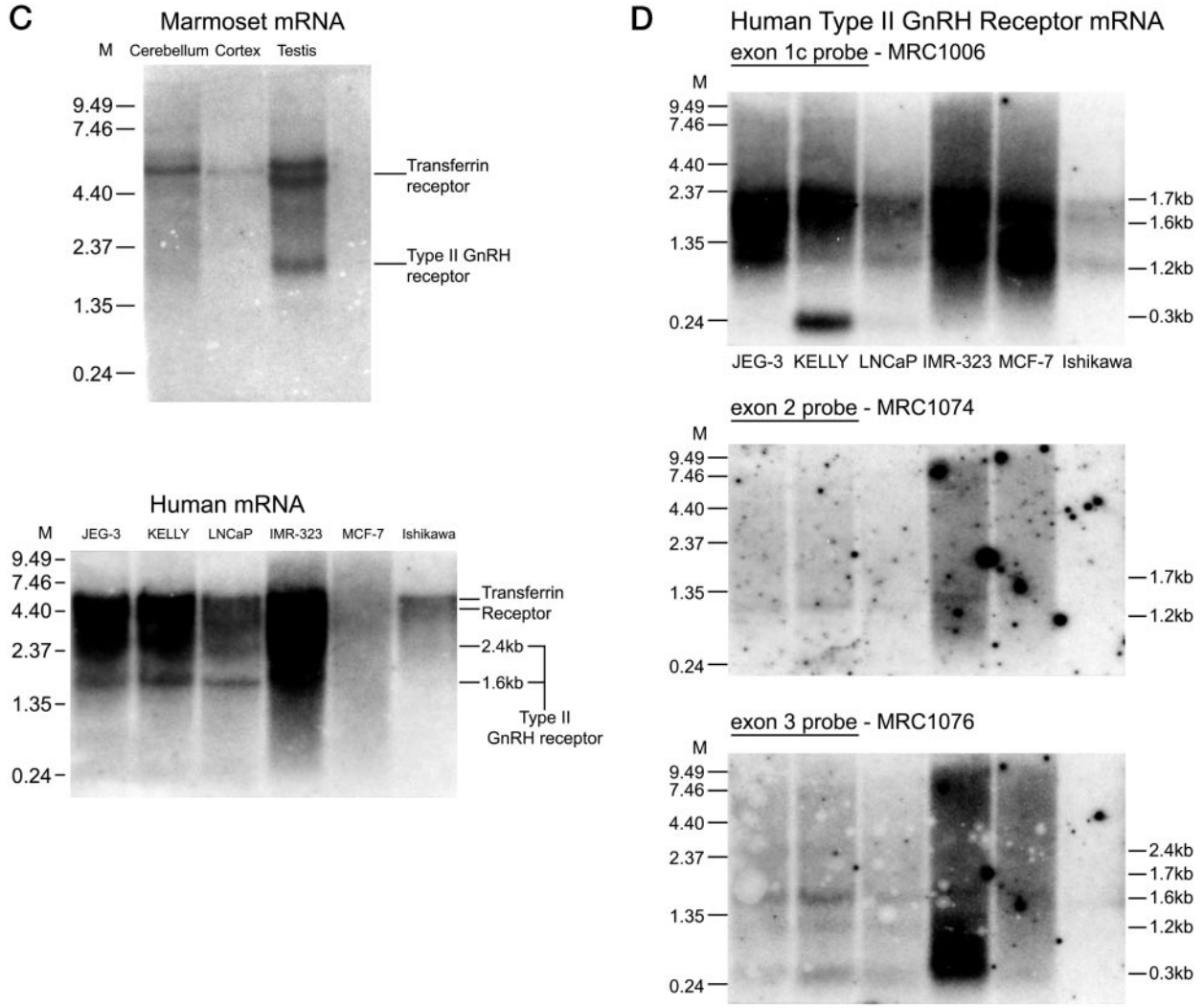

FIG. 2. *Continued*

Arg179 in the marmoset polypeptide to an in-frame UGA premature stop codon in human type II GnRH receptor exon 2 (Figs. 1, A and D–F; and 3A).

Phylogenetic analyses

Analysis of the nonhuman EST, nonmouse EST, and HTG GenBank databases revealed sequence entries corresponding to a type II GnRH receptor exon 2 for chimpanzee (GenBank accession no. AG122659), pig (GenBank accession no. BF702918), and cow (GenBank accession no. AV604131). These sequences encompass exon 2 and include short stretches of intron sequence (Fig. 3A). DNA sequence alignments indicate interspecies variations associated with occurrence of the premature stop codon. The chimpanzee sequence contains a positionally conserved UAA stop codon, whereas the porcine sequence contains a CGA (Arg) codon like the monkey type II GnRH receptor genes. Perhaps surprisingly, the bovine sequence contains a UGA stop codon shifted 3 bp to the 5′ direction. Phylogenetic tree construction, using the predicted amino acid sequences for ECL-2 and TMD-V, is problematic because of the short length and interrupted nature (*i.e.* the stop codons) of the data. It is reasonable to hypothesize that the primate sequences are more closely related to the porcine sequence, whereas the bovine sequence has probably been altered since artiodactyl divergence (Fig. 3, B and C).

Alternative splicing can partially compensate for gene disruption

A BLAST search of the GenBank human EST database revealed multiple entries spanning different fragments of the gene. These EST entries are derived from many different tissue types, but none of them represent sequence derived from a fully spliced transcript encoding a receptor homologous to the entire length of the 7-TMD marmoset type II GnRH receptor.

A human EST derived from a prostate adenocarcinoma cell line (GenBank accession no. BG036291) exhibits features consistent with alternative splicing within the first coding exon, such that the frameshift is avoided. A 260-bp sequence, flanking TMD I and including most of TMD II, is removed from this transcript. Though this EST exhibits splicing at the

TABLE 1. Organisation of overlapping genes on chromosome 1q.12

Accession no.	Nucleotide location	Description
Chromosome 1q.12		
AL160282	77,167–76,614 (555 bp)	Type II GnRH receptor coding exon 1
		Start ATG at 77,122
	75,640–76,191 (554 bp)	Intron 1 LINE 1.3—like fragment
	72,367–72,153 (215 bp)	Exon 2
		Premature stop codon at 72,339
	71,704–70,993 (716 bp)	Exon 3
		Stop codon at 71,287
	69,072–69,158 (87 bp)	RBM8A coding exon 1
		Start ATG at 69,093
	69,440–69,500 (61 bp)	Exon 2
	69,634–69,709 (76 bp)	Exon 3
	69,900–70,037 (138 bp)	Exon 4
	70,339–70,477 (139 bp)	Exon 5
	70,589–70,794 (206 bp)	Exon 6
		Stop codon at 70,634
	77,808–77,882 (75 bp)	Pex11β coding exon 1
		Start ATG at 77,825
	78,695–78,813 (119 bp)	Exon 2
	79,496–79,699 (204 bp)	Exon 3
	83,935–84,325 (1222 bp)	Exon 4
		Stop codon at 84,342

TABLE 2. Regions of sequence homology on chromosomes 1, 11, and 14

Accession no.	Nucleotide location	Description
Chromosome 1q.12		
AL160282	78,217–71,245 (6973 bp)	Matches AF403014 (chromosome 1)
	72,868–70,589 (2281 bp)	Matches AL132778 (chromosome 14)
	69,438–68,826 (615 bp)	Matches AC107240 (chromosome 11)

expected exon 1 splice donor site following the TMD IV coding region, it ends by splicing back into intron 1 at a splice acceptor site present in the partial LINE 1.3-like fragment (Fig. 1D). The encoded protein is shown in Figs. 1F and 2A. The DNA sequence at the end of this EST diverges from the chromosome 1 sequence, perhaps because of poor sequence quality, but it does extend to within the vicinity of a polyadenylation signal motif (AATAAA). Three more AATAAA sequences are present further downstream in intron 1 (Fig. 1A).

A misleading human EST (GenBank accession no. BG030969, Fig. 1D) can be mistaken to indicate the presence of a cryptic splice-acceptor located one base downstream from the UGA premature stop codon. However, sequence alignment with the chromosome 1q.12 clone indicates that this EST results from poor automated sequence quality over a region of 188 bp at the 3′ end of the EST (corresponding to the region 5′ upstream from the UGA premature stop codon, line marked in *bold* in Fig. 1D). In addition, this EST retains intron 2.

A selenocysteine incorporation sequence (SECIS)-like stem-loop in the 3′UTR of the type II GnRH receptor gene

Certain premature stop codons are translated as selenocysteine. RNA-fold analysis of the human type II GnRH receptor gene 3′UTR reveals a SECIS-like stem-loop structure (47, 48) 260 bp downstream from the exon 3 stop codon (Fig. 1 G). This structure encompasses 51 bases and possesses 4 non-Crick-Watson base pairs seen in SECIS elements. However, the expected stem structure following the 4-bp motif is absent in this case, and the conserved AAG/A triplet occupies a position 3′ to the stem-loop, unlike those seen in consensus SECIS elements. The SECIS-like sequence is followed by 2 direct repeat sequences of $(A)_8G$, located 48 bp downstream, perhaps indicative of the evolutionary deletion of a repetitive sequence element from between the $(A)_8G$ repeats (Fig. 1A).

Analysis of gene expression

We used different combinations of PCR primers and Northern blotting to analyze the pattern of gene expression of the human type II GnRH receptor-like sequence in a number of human cell lines and compared the results to type II GnRH receptor expression in marmoset tissues.

RT-PCR analyses of human type II GnRH receptor gene expression

It is possible to amplify a cDNA fragment derived from the whole of the human type II GnRH receptor gene on chromosome 1q.12 using RT-PCR primers MRC862 and MRC869 with total RNA isolated from HP-75 (pituitary adenoma) and IMR-32 (neuroblastoma) cell lines. Southern blotting and sequential rounds of DNA probe hybridization showed that a 1.17-kb cDNA product (corresponding to full-length PCR amplicon) and a 0.912-kb cDNA product (attributable to a deletion of 260 bp from the PCR amplicon) from these cell lines both hybridize to probes located in exons 1, 2, and 3 (Fig. 2, A and B). Demonstration of receptor gene expression in other human cell lines was hampered by high background

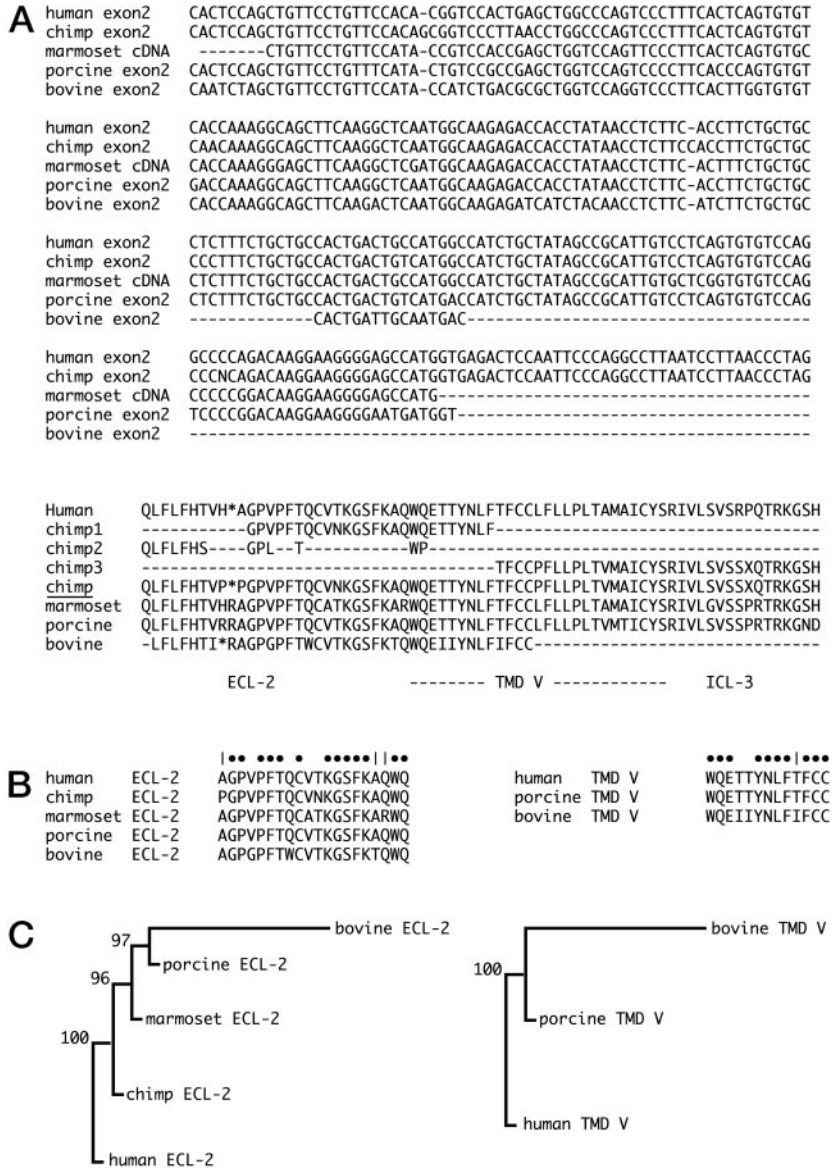

FIG. 3. A, Alignment of type II GnRH receptor gene exon 2 sequences from different species. B, Alignment of predicted peptide sequences from A produced by correcting two presumed DNA sequencing errors in the chimpanzee fragment The corrected sequence (chimp) is aligned below the raw data translations. C, Phylogenetic trees constructed from peptide sequences from B using PHYLIP software. High bootstrap probability values indicate the strength of support for the relationship. Branch lengths are proportional to evolutionary rate. Note the limitation that human and porcine sequences used are identical to each other in these analyses.

amplification, perhaps caused by interference by RBM8A transcripts. Therefore, the PCR results with this primer pair confirm the existence of alternative splicing of transcripts derived from this gene in some cell types.

Furthermore, we cloned and sequenced, from the HP-75 cell RT-PCR products, a full-length cDNA fragment (containing spliced exons 1–3) with frameshift and premature stop codon and a cDNA fragment possessing the 260-bp in-frame deletion from within exon 1 identical to the deletion seen in EST BG036291. This alternatively spliced fragment is labeled exon 1b in Fig. 2A. The HP-75 cell cDNA clone also contained exons 2 and 3, spliced together as expected, with a premature stop codon in the ECL-2 coding region. So far, there is no evidence that the stop codon is altered by post-transcriptional editing.

For a more systematic RT-PCR analysis of expression and splicing patterns from the type II GnRH receptor gene, we designed primers to subregions of the gene (Fig. 2A) and performed further amplifications.

Primers MRC862 and MRC924 generated an unexpected band at 0.9 kb and the expected bands at 0.686 kb and 0.427 kb, the latter being consistent with splicing out of exon 1b. Likewise, primers MRC862 and MRC1084 or MRC862 and MRC1085 also generated PCR products of the expected sizes and smaller bands consistent with removal of exon 1b.

Primers MRC1001 and MRC869 generated a 1.097-kb PCR product, and MRC1001 and MRC924 generated a 0.613-kb product, as expected (Fig. 2B). Both reactions also generated unexpected smaller bands.

Interestingly, primers MRC923 and MRC869 generated the expected 0.822-kb product and an unexpected product of 0.55 kb. Primers MRC923 and MRC924 generated a band of the expected size, 0.338 kb, and an unexpected band of less than 0.25 kb. These results suggest the potential for alternative splicing in the region spanning the premature stop codon.

Attempts to generate PCR products of the expected size,

using primers MRC1081 and MRC869 or primers MRC1083 and MRC869, failed.

Northern blotting

PolyA+ RNA, from a range of marmoset tissues and human cell lines, was probed to further investigate patterns of transcription from the type II GnRH receptor gene. It was not possible to isolate enough total RNA from HP-75 cells for polyA+ RNA preparation, because these cell commit to apoptosis within three to four passages *in vitro* (49), and their commercial availability has become problematic.

Probe hybridization indicated the existence of a single major transcript from the marmoset receptor gene in testis, in contrast to a number of different receptor transcripts in human cell lines (Fig. 2, C and D). The marmoset type II GnRH receptor mRNA is approximately 1.8-kb long. Strand-specific oligonucleotide probes detected bands of approximately 2.4 kb, 1.7 kb, 1.6 kb, 1.2 kb, and 0.3 kb in the human cell lines studied (Fig. 2D).

A 2.4-kb human gene transcript [the same size as that reported previously (5)] was detected in JEG-3, KELLY, LNCaP, and IMR-32 but not in MCF-7 or Ishikawa cell lines using a double-stranded DNA exon 1 probe (prepared using PCR primers MRC862 and MRC781) and with an oligonucleotide probe to exon 1 (MRC1006, Figs. 2, C and D).

A 1.7-kb band was detected with a probe to exon 1 (oligonucleotide MRC1006) in all of the cell lines studied (Fig. 2D).

A 1.6-kb band was detected with probes to exon 1 (DNA probe MRC862-MRC781 and oligonucleotide MRC1006) and exon 3 (oligonucleotide MRC1076) in JEG-3, KELLY, LNCaP, and Ishikawa cells (Fig. 2, C and D).

A 1.2-kb band was detected with probes to exon 1 (oligonucleotide MRC1006), exon 2 (oligonucleotide MRC1074), and exon 3 (oligonucleotide MRC1076) in JEG-3, KELLY, LNCaP, and IMR-32 cell lines (Fig. 2D).

A 0.3-kb band was detected with probes to exon 1 and exon 3, most noticeably in the KELLY neuroblastoma cell line (Fig. 2D).

A double-stranded DNA probe to the part of exon 1 removed by alternative splicing (produced using primers MRC1001-MRC1002) failed to detect any transcripts on the same Northern blot when stripped and reprobed.

To summarize, transcripts from exon 1 appear in more size classes than those from exon 2 or exon 3; and probes to exons 1, 2, and 3 all hybridize to a 1.2-kb transcript.

Sequences overlapping the type II GnRH receptor gene

Alignment of the RBM8A cDNA (GenBank accession no. AF231511) with the human chromosome 1q.12 HTG sequence reveals that the RBM8A gene encodes a 156-amino-acid protein in six exons (Fig. 1A) (25, 26). Polyadenylation signals for 5.5-kb, 3.2-kb, and 0.9-kb transcripts have been proposed previously (25, 27), although a total of eight possible AATAAA sites exist in the 3′UTR before the PEX11β gene. The 5′ flank of the RBM8A gene contains dense clusters of *Alu*-like fragments (Fig. 1A), some flanked by $(T)_nG$ motifs.

Alternative splicing in RBM8A transcripts

The codon at position Glu43 of RBM8A occupies a site involved in alternative splicing (25, 26). There are at least 50 human ESTs spanning the Glu43 coding region. The GAA codon (Glu43) is present in 86% of these. A Glu43-minus variant (Glu43−) arises from a transcript that uses a different splice acceptor site located three bases downstream from the GAA codon (25, 26).

Further characterization of the RBM8B locus

A HTG sequence on chromosome 14q.22-23 contains a spliced RBM8-coding sequence indicative of a processed pseudogene (25, 26) (GenBank accession no. AL132778 and AF231512), designated RBM8B (Fig. 1B).

Eight nucleotide mutations, leading to nonconservative substitutions, and one in-frame premature stop codon (TAA) are present within the RBM8B pseudogene (25, 26). Alignment of the chromosome 14q.22-23 RBM8B sequence with several hundred human ESTs revealed that none corresponded exactly to the chromosome 14q.22-23 sequence in this eight-base signature. This result suggests that RBM8B is transcriptionally inactive.

The 5′ flank of the RBM8B gene contains repetitive sequence fragments. A LINE 1.3-like fragment proximal to the coding region is followed by a 5′-distal *Alu* fragment and then a dense cluster of *Alu* fragments similar to those flanking the 5′ end of the RBM8A gene (Fig. 1, A and B).

Characterization of the PEX11β gene

BLAST search analysis of the GenBank databases, using the putative upstream promoter region for the human type II GnRH receptor gene, indicated the presence of the gene encoding the 259-amino-acid residue peroxisomal membrane protein pex11β (GenBank accession no. AB018080) (30). This gene lies in the opposite orientation to the receptor gene on the same chromosome 1q.12 HTG clone (GenBank accession no. AL160282). The exon-intron organization of the PEX11β gene was determined by alignment of PEX11β cDNA fragment sequences with the chromosome 1q.12 sequence. The PEX11β gene is arranged as 4 coding exons (Fig. 1A). Introns 2 and 3 contain *Alu*-like fragments, and the gene is flanked by *Alu*-like fragments in the 3′UTR. The first AATAAA sequence is located 1431 bp from the translation stop codon (UGA), and a second motif occurs 421 bp further downstream, suggesting that mature PEX11β mRNA occurs in the 2.3- to 2.7-kb size range.

The promoter element between the type II GnRH receptor and PEX11β genes

Further analyses of the EST database revealed the existence of a putative PEX11β gene transcription start site (GenBank accession no. BM010359) lying only 488 bp from a type II GnRH receptor gene transcription start site (GenBank accession no. BI850619, Fig. 1C). We screened the 695-bp DNA sequence lying between the receptor and PEX11β ATG translation start codons for consensus transcription factor binding sites. We found that potential binding sites were well separated on the opposite DNA strands, with the exception of

a shared myocyte enhancer factor 2 site (**TAAA**, Fig. 1C). The core promoter regions for each gene are separated by approximately 60 bp, and they each contain two potential SP-1 binding sites (**GGCG**). Both promoter regions lack obvious TATA elements, although sequences resembling muscle type TATA elements occur 259 bp upstream to the PEX11β translation start ATG. Significantly, a potential Pit-1 binding site (AAAT**ATTC**AT), upstream from the receptor gene, and a sequence similar to yeast UAS1 transcription factor binding site ($CTCCATN_5CTGGAG$), upstream from the PEX11β gene, may contribute to independent regulation of these genes (50).

Discussion

Type II GnRH may play an important role in reproductive physiology by eliciting specific cellular responses in different tissues (51–53). Genes encoding high-affinity GPCRs for type II GnRH have recently been characterized in some vertebrate species (5, 23, 24), and a functional human homolog might therefore be expected (4). Millar *et al.* (5) have presented mRNA and immunostaining evidence for type II GnRH receptor-like expression in human tissues, and Neill *et al.* (24, 53) have detected receptor-like RNA using human cDNA microarrays.

We have extended the recent analysis of homologous regions of human chromosomes 1q.12 and 14q.22-23 (25–29) and also identified another short region of homology to the RBM8A gene on chromosome 11. This study shows that the putative type II GnRH receptor gene on chromosome 1q.12 not only overlaps the RBM8A gene in the antisense orientation (25–29) but lies in close proximity to the PEX11β gene, which is also encoded in the antisense direction.

The human type II GnRH receptor gene homolog is apparently disrupted by a frameshift and a premature stop codon, and its expression may be subject to transcript-interference from RBM8A RNA (25, 26), as reflected by the RT-PCR difficulties encountered in this study. Although the 5′ promoter regions of the receptor and PEX11β genes overlap, their potential transcription factor binding sites remain distinct, as described for other genes arranged in head-to-head orientation (54, 55), suggesting that the type II GnRH receptor gene promoter may retain the capacity for tissue-specific regulation.

Here, we have demonstrated that the human type II GnRH receptor gene homolog is transcriptionally active in a variety of cell types and that a number of different transcripts exist. In fact, multiple RNA transcript species generated by alternative splicing are a common feature of GnRH receptor genes (33, 34) and related GPCR genes (31–32). Although full-length type II GnRH receptor transcripts are present in marmoset, rhesus, and green monkey, a full-length transcript encoding an uninterrupted 7-TMD human type II GnRH receptor ORF has not yet been isolated. RT-PCR-Southern blotting and cDNA sequencing data indicated that transcripts containing exons 1, 2, and 3 do occur, but the frameshift and premature stop codon seem not to be altered by posttranscriptional RNA editing. These results are consistent with a recently described transcript identified in human testis tissue (GenBank accession no. AY081843), where the frameshift is circumvented by removal of exon 1b through alternative splicing.

Neill *et al.* (24, 27) detected widespread type II GnRH receptor-like gene expression in human tissues, using a full-length riboprobe to screen a cDNA microarray. However, they did not fully determine the riboprobe specificity by cloning cDNAs hybridizing to this probe. Interpretation of the widespread expression is also dependent on the subtraction of any nonspecific background signal from the array. Further analyses of receptor gene expression in human tissues are required.

A 2.4-kb receptor gene transcript is apparent in mRNA from human tissues (5) and cell lines when using a double-stranded exon 1 DNA probe in Northern blotting. This agrees with the predicted approximate size for a full-length type II GnRH receptor transcript containing a long 3′UTR (1224 bases from exon 3 UGA to the nearest AATAAA, in the absence of 3′UTR splicing). However, this transcript is similar in size to that expected for PEX11β mRNA; and, since the transcription start sites for the PEX11β gene have not been determined, a double-stranded probe might not distinguish between transcripts from the two genes. The Northern blots of Millar *et al.* (5) may therefore be misleading, depending on the level of PEX 11β expression in different tissues. For this reason, and because RBM8A transcripts of 3.2, 5.5, and 7.5 kb are so abundantly expressed (25), sequential hybridization studies using single-stranded probes on the same Northern blot filter were undertaken. In this way, a variety of receptor-strand transcripts smaller than 2.4 kb were identified, some of which are expressed in a cell-type-specific fashion. The detection of several bands after Northern blotting is consistent with the RT-PCR evidence that alternative processing of transcripts occurs in the HP-75, IMR-32, and LNCaP cell lines. The results of Northern blotting are difficult to interpret fully, because of the complexity of the hybridization signals and because the sites of transcription initiation and 5′ or 3′UTR processing are currently unknown. Protein coding exons 1, 2, and 3 can be detected in a transcript of approximately 1.2 kb. In addition, transcripts containing exon 1 exist in several size classes, compared with those from exon 2 or exon 3. This suggests frequent splicing from exon 1 into intron 1 and utilization of the polyadenylation signals located nearby (discussed below). However, the DNA sequence identities of all possible RT-PCR products derived from type II GnRH receptor or RBM8A gene expression remain to be determined, and further studies should enable a complete characterization of their transcription patterns. Similar analyses are required on human cell lines in which type II GnRH receptor gene expression has been partially characterized by RT-PCR. The data of Grundker *et al.* (56) shows mRNA in which exons 1 and 2 are spliced together but does not address alternative splicing. Northern blotting or further RT-PCR studies are required to assess the number of transcripts in these cells.

Although altered gene function, as a consequence of reading frame disruption, has been acknowledged previously, the phenomenon has been explored in relatively little detail for transcriptionally active GPCR genes with disrupted ORFs. Millar *et al.* (5) presented immunocytochemical evidence that the ECL-3 epitope, encoded downstream from the

premature stop codon, may be detectable in human pituitary. This observation, and the fact that a cognate receptor for the human GnRH-II ligand might be expected to exist, prompts a discussion of possible translation pathways. Biosynthesis of a conventionally functional human type II GnRH receptor with the classical 7-TMD helix structure would require that two translational recoding mechanisms be invoked.

First, the −1 frameshift within exon 1 must be corrected or translation initiation must occur at an alternative downstream translation initiation site. However, we find only weak matches, with 60% identity to the Kozak consensus sequences, in the downstream vicinity of the frameshift, and there is no evidence for an alternative exon 1 on chromosome 1. Examples of eukaryotic cellular mRNAs that use programmed −1 ribosomal frameshifting to control protein expression have been discovered recently (57, 58), but the −1 frameshift site in the type II GnRH receptor gene (**G GG**T CAG) does not contain the necessary heptamer "slippery site" consensus sequence (**X XX**Y YYZ, where **X XX** are any three identical nucleotides constituting the start of the –1 reading frame, YYY is either AAA or UUU, and Z is either A, U, or C). Thus, there is currently no evidence for an alternative site of translation initiation in exon 1.

Part of exon 1 (exon 1b) is frequently excised from transcripts in which exons 1, 2, and 3 are spliced together. The putative methionine initiation codon within this transcript is located in a region (GCCACC**ATG**T) with 90% identity to the 10-base core-consensus Kozak translation initiation element (GCCA/GCC**ATG**G) (39). This cDNA bypasses the frameshift and encodes a polypeptide with two TMDs before the premature stop codon, similar to GenBank accession no. AY081843 and EST BG036291. The latter EST exhibits alternative transcript processing at its 3′ end, through utilization of a splice acceptor located within the LINE 1.3-like fragment in intron 1. This transcript probably becomes polyadenylated because of the AATAAA signal motifs present at the 3′ flank of the repetitive sequence element (Fig. 1, A and D).

Second, the in-frame UGA stop codon within exon 2 must be either skipped or translated differently. Decoding of a UGA stop codon as selenocysteine requires the presence of a selenocysteine insertion element (SECIS) in the 3′UTR of the mRNA (47, 48). A scan of the putative 3′UTR of the type II GnRH receptor gene for the SECIS pattern indicated a potential stem-loop structure with some similarity to current models of SECIS structure. However, incorporation of ^{75}Se-selenocysteine into protein in COS cells transfected with human type II GnRH receptor cDNA did not reveal a detectable species in the molecular weight range consistent with 292 (5-TMD)- or 379 (7-TMD)-amino-acid residue receptors (unpublished data).

An alternative hypothesis is that a protein is translated downstream from the premature UGA stop codon. Such an arrangement may explain why the position of the UGA premature stop codon has been selected independently on at least two occasions during evolution (in certain primates and artiodactyls). There is a candidate translation initiation site downstream from the premature stop codon (ACTGCC**ATG**G in human and ACTGTC**ATG**G in chimpanzee, Figs. 1E and 3A), with 80% identity to the Kozak consensus sequence (39). Translation from this site would generate a protein with at least two TMDs (equivalent to TMDs VI and VII of the monkey receptors), ECL-3 and the cytoplasmic tail. However, the functional activity of Kozak-like translation initiation sequences in this region of the gene have yet to be tested. Alternative translation initiation, or selenocysteine incorporation into full-length receptor in certain cell types, may therefore account for immunocytochemical detection of ECL-3 epitope in human pituitary (5).

It may be possible that partial-receptor gene products, described in *Results* (Figs. 1F and 2A), are capable of interaction with the type I GnRH receptor or with other GPCRs, through lateral interaction or domain-swapping, as hypothesized for some receptors (59, 60). Among GnRH receptors, expression of alternatively spliced transcripts has been shown to influence the function of full-length type I GnRH receptor (33, 34). Alternatively, they might assemble to form a 5-TMD type receptor described in other GPCR systems (61). Perhaps unusual cross-species variations in the organization and evolution of GnRH receptor-derived molecules are not entirely unexpected. The evolutionary need for diversity in the reproductive biology of animal species may have exerted pressure to adapt the type II GnRH receptor protein to perform subtly different functions.

Considering the evolution of the human type II GnRH receptor gene further, we note that deletion of the G residue in exon 1 could have arisen as a consequence of DNA replication arrest at a frameshift hotspot (T G A/G A/G G/T A/C) (62). Alternatively, the frameshift may have arisen by a DNA recombination event between an ancestral form of the receptor gene and a reverse-transcribed receptor transcript.

Hypothetically, RT could have been primed by RNA duplex formation with an RBM8A transcript or, alternatively, by self-priming from the LINE 1.3-like fragment (63, 64) or from either of the oligo(deoxythymidine) sequences in intron 1.

The premature stop codon (UGA) in the human and bovine receptor genes may have arisen independently from two different CGA (Arg) codons in an ancestral vertebrate gene. Independent evolutionary disruption events have been described for other GPCR genes (65, 66). Demethylation of methyl cytosine to produce uracil with subsequent DNA repair, generating a C-to-T transition, is recognized as a frequent event in the evolution of genes (67). The premature stop codon (UAA) in the chimpanzee receptor gene fragment most likely originated from an ancestral CAA (Gln) codon, because a Gln residue at this position in ECL-2 would represent a conservative amino acid change, relative to the Arg residue present in the monkey and porcine receptors. Indeed, the interchange of Gln and Arg residues is the commonest conservative change observed in evolution.

Alu-element mediated methylation of the C residues in adjacent DNA sequences is thought to promote the likelihood of C-to-T transitions in genes subject to germline patterns of expression in higher vertebrates (68, 69), although all the *Alu*-like sequences at the human type II GnRH receptor gene locus are fragmentary and likely to be inactive. A differential distribution of *Alu*-like and LINE-like elements in the marmoset type II GnRH receptor gene may help to explain why the gene is not disrupted in this species.

In conclusion, our observations indicate that the human

type II GnRH receptor gene is expressed as a variety of splice variants, and further investigations are required to determine whether these translate to functional proteins in man and in other mammals possessing type II GnRH receptor gene homologs with disrupted ORFs.

Acknowledgments

We would like to thank Pamela Brown for critically reading the manuscript, Stuart Maudsley and Lindsay Davidson for useful discussions, and Nicola Miller for assistance with cell culture. Thanks also to Eleanor Meikle and Ted Pinner for graphics.

Received June 13, 2002. Accepted October 16, 2002.

Address all correspondence and requests for reprints to: Kevin Morgan, Medical Research Council Human Reproductive Sciences Unit (K.M., A.J.P., R.S., T.R.O., R.P.M.), University of Edinburgh Academic Centre, 49 Little France Crescent, Old Dalkeith Road, Edinburgh EH16 4SB, United Kingdom. E-mail: k.morgan@hrsu.mrc.ac.uk.

This work was supported by Grant 060257 (to A.J.P.) from The Wellcome Trust.

GenBank accession no. for this manuscript: ALIGN_000364.

References

1. **King JA, Millar RP** 1979 Heterogeneity of vertebrate luteinizing hormone releasing hormone. Science 206:67–69
2. **Sherwood NM, Lovejoy DA, Coe IR** 1993 Origin of mammalian gonadotropin-releasing hormones. Endocr Rev 14:241–254
3. **Sealfon SC, Weinstein H, Millar RP** 1997 Molecular mechanisms of ligand interaction with the gonadotropin-releasing hormone receptor. Endocr Rev 18:180–205
4. **White RB, Eisen JA, Kasten TL, Fernald RD** 1998 Second gene for gonadotropin-releasing hormone in humans. Proc Natl Acad Sci USA 95:305–309
5. **Millar R, Lowe S, Conklin D, Pawson A, Maudsley S, Troskie B, Ott T, Millar M, Lincoln G, Sellar R, Faurholm B, Scobie G, Kuestner R, Terasawa E, Katz A** 2001 A novel mammalian receptor for the evolutionarily conserved type II gonadotropin releasing hormone. Proc Natl Acad Sci USA 98:9636–9641
6. **Zoeller RT, Seeburg PH, Young 3rd WS** 1988 *In situ* hybridization histochemistry for messenger ribonucleic acid (mRNA) encoding gonadotropin-releasing hormone (GnRH): effect of estrogen on cellular levels of GnRH mRNA in female rat brain. Endocrinology 122:2570–2577
7. **Berry SA, Pescovitz OH** 1988 Identification of a rat GnRH-like substance and its messenger RNA in rat testis. Endocrinology 123:661–663
8. **Urbanski HF, White RB, Fernald RD, Kohama SG, Garyfallou VT, Densmore VS** 1999 Regional expression of mRNA encoding a second form of gonadotropin-releasing hormone in the macaque brain. Endocrinology 140:1945–1948
9. **Latimer VF, Kohama SG, Garyfallou VT, Urbanski HF** 2001 A developmental increase in the expression of messenger ribonucleic acid encoding a second form of gonadotropin-releasing hormone in the rhesus macaque hypothalamus. J Clin Endocrinol Metab 86:324–329
10. **Hamalainen T, Poutanen TM, Huhtaniemi I** 1999 Age- and sex-specific function of a 2-kilobase 5′-flanking sequence of the murine luteinizing hormone receptor gene in transgenic mice. Endocrinology 140:5322–5329
11. **Sakuma Y, Pfaff DW** 1980 LH-RH in the mesencephalic central grey can potentiate lordosis reflex of female rats. Nature 283:566–567
12. **Huirne JA, Lambalk CB** 2001 Gonadotropin-releasing-hormone-receptor antagonists. Lancet 358:1793–803
13. **Klein KO, Barnes KM, Jones JV, Feuillan PP, Cutler Jr GB** 2001 Increased final height in precocious puberty after long-term treatment with LHRH agonists: the National Institutes of Health experience. J Clin Endocrinol Metab 86:4711–4716
14. **Lazar L, Pertzelan A, Weintrob N, Phillip M, Kauli R** 2001 Sexual precocity in boys: accelerated versus slowly progressive puberty gonadotropin suppressive therapy and final height. J Clin Endocrinol Metab 86:4127–4133
15. **Windle JJ, Weiner RJ, Mellon PL** 1990 Cell lines of the pituitary gonadotrope lineage derived by targeted oncogenesis in transgenic mice. Mol Endocrinol 4:597–603
16. **Chen A, Laskar-Levy O, Ben-Aroya N, Koch Y** 2001 Transcriptional regulation of the human GnRH II gene is mediated by a putative cAMP response element. Endocrinology 142:3483–3492
17. **Limonta P, Moretti RM, Marelli MM, Dondi D, Parenti M, Motta M** 1999 The luteinizing hormone-releasing hormone receptor in human prostate cancer cells: messenger ribonucleic acid expression, molecular size, and signal transduction pathway. Endocrinology 140:5250–5256
18. **Cheng KW, Chow BKC, Leung PCK** 2001 Functional mapping of a placenta-specific upstream promoter for human gonadotropin-releasing hormone receptor gene. Endocrinology 142:1506–1516
19. **Beranova M, Oliveira LMB, Bedecarrats GY, Schipani E, Vallejo M, Ammini AC, Quintos JB, Hall JE, Martin KA, Hayes FJ, Pitteloud N, Kaiser UB, Crowley Jr WF, Seminara SB** 2001 Prevalence, phenotypic spectrum, and modes of inheritance of gonadotropin-releasing hormone receptor mutations in idiopathic hypogonadotropic hypogonadism. J Clin Endocrinol Metab 86:1580–1588
20. **Schally AV** 1999 Luteinizing hormone-releasing hormone analogs: their impact on the control of tumorigenesis. Peptides 20:1247–1262
21. **Schally AV, Comaru-Schally AM, Plonowski A, Nagy A, Halmos G, Rekasi Z** 2000 Peptide analogs in the therapy of prostate cancer. Prostate 45:158–166
22. **Jonat W** 2001 Goserelin (Zoladex)—its role in early breast cancer in pre- and perimenopausal women. Br J Cancer 85(Suppl 2):1–5
23. **Wang L, Bogerd J, Choi HS, Seong JY, Soh JM, Chun SY, Blomenrohr M, Troskie BE, Millar RP, Yu WH, McCann SM, Kwon HB** 2001 Three distinct types of GnRH receptor characterized in the bullfrog. Proc Natl Acad Sci USA 98:361–366
24. **Neill JD, Duck W, Sellers JC, Musgrove LC** 2001 A gonadotropin-releasing hormone (GnRH) receptor specific for GnRH II in primates. Biochem Biophys Res Commun 282:1012–1018
25. **Millar R, Conklin D, Lofton-Day C, Hutchinson E, Troskie B, Illing N, Sealfon SC, Hapgood J** 1999 A novel human GnRH receptor homolog gene: abundant and wide tissue distribution of the antisense transcript. J Endocrinol 162:117–126
26. **Faurholm B, Millar RP, Katz AA** 2001 The genes encoding the type II gonadotropin-releasing hormone receptor and the ribonucleoprotein RBM8A in humans overlap in two genomic loci. Genomics 78:15–18
27. **Neill JD** 2002 Minireview: GnRH and GnRH receptor genes in the human genome. Endocrinology 143:737–743
28. **Conklin D, Rixon M, Kuestner R, Maurer M, Whitmore T, Millar R** 2000 Cloning and gene expression of a novel human ribonucleoprotein. Biochim Biophys Acta 1492:465–469
29. **Salicioni AM, Xi M, Vanderveer LA, Balsara B, Testa JR, Dunbrack Jr RL, Godwin AK** 2000 Identification and structural analysis of human RBM8A and RBM8B: two highly conserved RNA-binding motif proteins that interact with OVCA1, a candidate tumor suppressor. Genomics 69:54–62
30. **Schrader M, Reuber BE, Morrell JC, Jimenez-Sanchez G, Obie C, Stroh TA, Valle D, Schroer TA, Gould SJ** 1998 Expression of *Pex11β* mediates peroxisome proliferation in the absence of extracellular stimuli. J Biol Chem 273:29606–29614
31. **Petersen S, Rasch AC, Penshorn M, Beil FU, Schulte HM** 2001 Genomic structure and transcriptional regulation of the human growth hormone secretagogue receptor. Endocrinology 142:2649–2659
32. **Rekasi Z, Czompoly T, Schally AV, Halmos G** 2000 Isolation and sequencing of cDNA splice variants of growth hormone-releasing hormone receptors from human cancers. Proc Natl Acad Sci USA 97:10561–10566
33. **Wang L, Oh DY, Bogerd J, Choi HS, Ahn RS, Seong JY, Kwon HB** 2001 Inhibitory activity of alternative splice variants of the bullfrog GnRH receptor-3 on wild-type receptor signalling. Endocrinology 142:4015–4025
34. **Grosse R, Schöneberg T, Shultz G, Gudermann T** 1997 Inhibition of gonadotropin-releasing hormone receptor signaling by expression of a splice variant of the human receptor. Mol Endocrinol 11:1305–1318
35. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ** 1990 Basic local alignment search tool. J Mol Biol 215:403–410
36. **Yee D, Conklin D,** Automated clustering and assembly of large EST collections. Proc Sixth International Conference on Intelligent Systems for Molecular Biology, Montréal, Canada, 1998, AAAI Press, pp 203–211
37. **Sonnhammer E, Durbin R** 1995 A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 167:1–10
38. **Quandt K, Frech K, Karas H, Wingender E, Werner T** 1995 MatInd and MatInspector—new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res 23:4878–4884
39. **Kozak M** 1999 Initiation of translation in prokaryotes and eukaryotes. Gene 234:187–208
40. **Rost B, Fariselli P, Casadio R** 1996 Topology prediction for helical transmembrane proteins at 86% accuracy. Protein Sci 5:1704–1718
41. **Zuker M, Mathews DH, Turner DH** 1999 Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In: Barciszewski J, Clark BFC, eds. RNA biochemistry and biotechnology NATO ASI Series. Dordrecht, The Netherlands: Kluwer Academic Publishers; 11–43
42. **Mathews DH, Sabina J, Zuker M, Turner DH** 1999 Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J Mol Biol 288:911–940
43. **Retief JJ** 2000 Phylogenetic analysis using PHYLIP. Methods Mol Biol 132:243–25844
44. **Chomczynski P, Sacchi N** 1987 Single-step method of RNA isolation by acid guanidinium thiocyanate-acid phenol-chloroform extraction. Anal Biochem 162:156–159
45. **Sambrook J, Fritsch EF, Maniatis T** 1989 Molecular cloning. A laboratory manual. 2nd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 7.43–7.45
46. **Proudfoot N** 1991 Poly(A) signals. Cell 64:671–674

47. **Lescure A, Gautheret D, Carbon P, Krol A** 1999 Novel selenoproteins identified *in silico* and *in vivo* by using a conserved RNA structural motif. J Biol Chem 274:38147–38154
48. **Fagegaltier D, Lescure A, Walczak R, Carbon P, Krol A** 2000 Structural analysis of new local features in SECIS RNA hairpins. Nucleic Acids Res 28:2679–2689
49. **Oka H, Jin L, Kulig E, Scheithauer BW, Lloyd RV** 1999 Pituitary adenylate cyclase-activating polypeptide inhibits transforming growth factor-beta 1-induced apoptosis in a human pituitary adenoma cell line. Am J Pathol 155:1893–1900
50. **Gurvitz A, Hiltunen JK, Erdmann R, Hamilton B, Hartig A, Ruis H, Rottensteiner H** 2001 Saccharomyces cerevisiae Adr1p governs fatty acid beta-oxidation and peroxisome proliferation by regulating POX1 and PEX11. J Biol Chem 276:31825–31830
51. **Maney DL, Richardson RD, Wingfield JC** 1997 Central administration of chicken gonadotropin-releasing hormone-II enhances courtship behaviour in a female sparrow. Horm Behav 32:11–18
52. **Lumpkin MD, Moltz JH, Yu WH, Samson WK, McCann SM** 1987 Purification of FSH-releasing factor: its dissimilarity from LHRH of mammalian, avian, and piscian origin. Brain Res Bull 18:175–178
53. **Neill JD** 2002 Mammalian gonadotropin-releasing hormone (GnRH) receptor subtypes. Arch Physiol Biochem 220:129–136
54. **Igaki H, Nakagawa K, Aoki Y, Ohtomo K, Kukimoto I, Kanda T** 2001 Characterization of the bi-directional transcriptional control region between the human UFD1L and CDC45L genes. Biochem Biophys Res Commun 283: 569–576
55. **Ito E, Xie G, Maruyama K, Palmer PP** 2000 A core-promoter region functions bi-directionally for human opioid-receptor-like gene ORL1 and its 5′-adjacent gene GAIP. J Mol Biol 304:259–270
56. **Grundker C, Gunthert AR, Millar RP, Emons G** 2002 Expression of gonadotropin-releasing hormone II (GnRH-II) receptor in human endometrial and ovarian cancer cells and effects of GnRH-II on tumor cell proliferation. J Clin Endocrinol Metab 87:1427–1430
57. **Shigemoto K, Brennan J, Walls E, Watson CJ, Stott D, Rigby PWJ, Reith AD** 2001 Identification and characterisation of a developmentally regulated mammalian gene that utilises –1 programmed ribosomal frameshifting. Nucleic Acids Res 29:4079–4088
58. **Hammell A, Taylor R, Peltz S, Dinman J** 1999 Identification of putative programmed ribosomal frameshift signals in large DNA databases. Genome Res 9:417–427
59. **Schultz A, Grosse R, Schultz G, Gudermann T, Schöneberg T** 2000 Structural implication for receptor oligomerization from functional reconstitution studies of mutant V2 vasopressin receptors. J Biol Chem 275:2381–2389
60. **Gouldson PR, Snell CR, Bywater RP, Higgs C, Reynolds CA** 1998 Domain swapping in G-coupled receptor dimers. Protein Eng 11:1181–1193
61. **Ling K, Wang P, Zhao J, Wu YL, Cheng ZJ, Wu GX, Hu W, Ma L, Pei G** 1999 Five-transmembrane domains appear sufficient for a G protein-coupled receptor: functional five-transmembrane domain chemokine receptors. Proc Natl Acad Sci USA 96:7922–7927
62. **Krawczak M, Cooper DN** 1991 Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. Hum Genet 86:425–441
63. **Esnault C, Maestre J, Heidmann T** 2000 Human LINE retrotransposons generate processed pseudogenes. Nat Genet 24:363–367
64. **Feng Q, Moran JV, Kazazian Jr HH, Boeke JD** 1996 Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87:905–916
65. **Matsumoto M, Nomura T, Momose K, Ikeda Y, Kondou Y, Akiho H, Togami J, Kimura Y, Okada M, Yamaguchi T** 1996 Inactivation of a novel neuropeptide Y/peptide YY receptor gene in primate species. J Biol Chem 271:27217–27220
66. **Starback P, Wraith A, Eriksson H, Larhammar D** 2000 Neuropeptide Y receptor gene y6: multiple deaths or resurrections? Biochem Biophys Res Commun 277:264–269
67. **Sved J, Bird A** 1990 The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. Proc Natl Acad Sci USA 87: 4692–4696
68. **Liu WM, Schmid CW** 1993 Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. Nucleic Acids Res 21: 1351–1359
69. **Schmid CW** 1998 Does SINE evolution preclude Alu function? Nucleic Acids Res 26:4541–4550