

# A novel human GnRH receptor homolog gene: abundant and wide tissue distribution of the antisense transcript

R Millar<sup>1</sup>, D Conklin<sup>2</sup>, C Lofton-Day<sup>2</sup>, E Hutchinson<sup>1</sup>, B Troskie<sup>1</sup>, N Illing<sup>1</sup>, S C Sealson<sup>3</sup> and J Hapgood<sup>1</sup>

<sup>1</sup>MRC Molecular Reproductive Endocrinology Research Unit, University of Cape Town Medical School, Observatory 7925, South Africa

<sup>2</sup>Zymogenetics, 1201 Eastlake Ave East, Seattle, Washington 98102, USA

<sup>3</sup>Fishberg Neurobiology Center and Department of Neurology, Mt Sinai School of Medicine, New York, USA

(Requests for offprints should be addressed to R Millar who is now at MRC Reproductive Biology Unit, Centre for Reproductive Biology, 37 Chalmers Street, Edinburgh EH3 9EW, UK)

(N Illing is now at Department of Biochemistry, University of Cape Town, South Africa)

(J Hapgood is now at Department of Biochemistry, University of Stellenbosch, Stellenbosch, South Africa)

## Abstract

Gonadotropin releasing hormone (GnRH) regulates the reproductive system through a specific G-protein-coupled receptor (GPCR) in pituitary gonadotropes. The existence of two (or more) forms of GnRH in most vertebrates suggested the existence of GnRH receptor subtypes (I and II). Using sequence information for extracellular loop 3 of a putative Type II GnRH receptor from a reptile species, we have looked for a Type II GnRH receptor gene in the human genome EST (expressed sequence tag) database. A homolog was identified which has 45% and 41% amino acid identity with exons 2 and 3 of the known human GnRH pituitary receptor (designated Type I) and much lower homology with all other GPCRs. A total of 27 contiguous ESTs was found and comprised a continuous sequence of 1642 nucleotides. The EST sequences were confirmed in the cloned human gene and in PCR products of cDNA from several tissues. All EST transcripts detected were in the antisense orientation with respect to the novel GnRH receptor sequence and were highly expressed in a wide range of human brain and peripheral

tissues. PCR of cDNA from a wide range of tissues revealed that intronic sequence equivalent to intron 2 of the Type I GnRH receptor was retained. The failure to splice out putative intron sequences in transcripts which spanned exon-intron boundaries is expected in antisense transcripts, as candidate donor and acceptor sites were only present in the gene when transcribed in the orientation encoding the GnRH receptor homolog. No transcripts extended 5' to the sequence corresponding to intron 2 of the Type I GnRH as the antisense transcripts terminated in poly A due to the presence of a polyadenylation signal sequence in the putative intron 2 when transcribed in the antisense orientation. These findings suggest that a Type II GnRH receptor gene has arisen during vertebrate evolution and is also present in the human. However, the receptor may have become vestigial in the human, possibly due to the abundant and universal tissue transcription of the opposite DNA strand to produce antisense RNA.

*Journal of Endocrinology* (1999) **162**, 117–126

## Introduction

The hypothalamic peptide, gonadotropin releasing hormone (GnRH) is the central regulator of the reproductive system in vertebrates. The demonstration of GnRH structural variants in vertebrates (King & Millar 1979) has led to the identification of eleven different structures (Millar & King 1988, Sherwood *et al.* 1993, King & Millar 1997, Millar *et al.* 1997, Sealson *et al.* 1997). A striking feature of the comparative evolution of the structures was the presence of two or more forms of GnRHs in most vertebrate species, suggesting an early gene duplication (Millar & King 1988). A particularly intriguing observation was the almost universal occurrence of a form of

GnRH first found in chicken brain and designated chicken GnRH II (Millar & King 1988). Since this form is present in most vertebrate species ranging from teleost fish to mammals (Millar & King 1988, Sherwood *et al.* 1993, King & Millar 1997, Millar *et al.* 1997, Sealson *et al.* 1997) it has clearly been structurally conserved over many millions of years and was postulated to have an important function, possibly as a neuromodulator in the central and peripheral nervous systems (Millar & King 1988). As the terminology 'chicken GnRH II' for a GnRH structure present in most vertebrates is likely to lead to confusion, we proposed that it be referred to as 'Type II GnRH' or 'GnRH II' (Millar *et al.* 1997, Sealson *et al.* 1997). The existence of several forms of the GnRH ligand within a

single species suggested that there may also be more than one form of receptor. To date only one form of GnRH receptor (designated here as Type I) has been cloned from mammals (Eidne *et al.* 1992, Kaiser *et al.* 1992, Kakar *et al.* 1992, Reinhart *et al.* 1992, Tsutsumi *et al.* 1992, Brooks *et al.* 1993, Chi *et al.* 1993, Illing *et al.* 1993, Kakar *et al.* 1993, Perrin *et al.* 1993, Troskie *et al.* 1998). We therefore designed a series of pairs of degenerate oligonucleotides to the mammalian GnRH receptor and cloned short sequences encoding the extracellular (EC) loop 3 domain (EC3) which suggested the presence of three different GnRH receptor genes in species of fish, amphibian, reptile and bird (Troskie *et al.* 1998). Two of these are most like the mammalian GnRH pituitary receptors (Type I) and were therefore designated Type Ia and Type Ib (Troskie *et al.* 1998). The third was more different and was designated Type II.

Taken together, these observations suggested that the early gene duplication of GnRH into a Type I and Type II ligand was accompanied by a gene duplication of the cognate receptors into a Type I and Type II receptor. We therefore searched the public human genome EST (expressed sequence tag) databases for a Type II homolog of the known pituitary GnRH receptor (Type I) and report here DNA sequences which encode protein sequence with about 42% amino acid identity with the pituitary Type I GnRH receptor in the region encompassing the extracellular loop 2 domain (EC2) through to the end of the transmembrane (TM) 7 domain.

## Materials and Methods

### *Search for GnRH receptor homologs*

A series of degenerate oligomer pairs had been designed for use in the PCR amplification of GnRH receptors in non-mammalian DNA. One pair was designed to amplify EC3 which is the domain shown to determine ligand specificity (Flanagan *et al.* 1994). With this pair we succeeded in obtaining the sequence of EC3 from fish, amphibian, reptile and bird species (Troskie *et al.* 1998). All of these species had a sequence homologous to the cloned mammalian GnRH receptors. However, in the amphibian and reptilian DNA distinctly different second putative GnRH receptors (designated Type II) with the EC3 sequences PEMLTEEKVPVLSHI and PEMLTRGQVPPSLSHI respectively, were found. The tblastn program (Altschul *et al.* 1990) was used to query the Merck-WashU EST database for homology to the reptile EC3 sequence and EST200 816 was identified (Troskie *et al.* 1998). The same approach was employed using the complete sequence of the human Type I GnRH receptor. Overlapping EST contigs were gathered and assembled using the Rex program which automatically gathers and assembles overlapping ESTs into longer 'virtual'

cDNA sequences (Yee & Conklin 1998). The blastx (Heywood 1986) and tfasta (Pearson *et al.* 1997) programs were used to further explore ESTs resulting from the above search.

### *PCR and sequence analysis of cDNA and genomic clones*

Oligonucleotide primers 10 063 (5'-CTGACCTTCAT CCTCTGCTGGACACC-3'), and 10 071 (5'-GGAGA GCAGGAGTAGAAGTGAG-3') were designed from the sequences of the two ESTs. Analysis of the EST sequences suggested that they represented a sequence corresponding to the EC3 of the Type I GnRH receptor through to the putative stop codon, and a 3' untranslated region of a putative human Type II GnRH receptor. The primers were used to obtain a 415 bp fragment that spans the region from the putative sixth transmembrane domain (TM6) through to the C-terminus of the Type I GnRH receptor. A panel of 'marathon ready' cDNA templates was prepared from human lymph node, placenta, uterus, liver, kidney, spleen, heart, skeletal muscle, ovary, testis, retina, thyroid, MCF 7 cell line, whole brain, cerebellum, hypothalamus, and fetal brain poly A RNA using a Marathon cDNA amplification kit (Clontech, Palo Alto, CA, USA) and oligo dt priming according to the protocol provided by the manufacturer. These templates were used in polymerase chain reactions to generate DNA encoding the human GnRH II receptor.

PCR amplification was carried out according to the manufacturer's instructions. Thirty picomoles of each primer were used in the reactions. Each PCR template was amplified for 35 cycles (94 °C, 20 s; 68 °C, 1 min) followed by a 10 min extension at 72 °C. PCR products were separated on 1.5% agarose gels. A 415 bp fragment was found in all cDNAs tested, with the strongest signal found in placenta, uterus and liver. The identity of the 415 bp fragment was confirmed by DNA sequencing.

The 415 bp fragment and primers 10 063 and 10 071 were used to screen a P1 genomic library (Genome Systems, St Louis, MO, USA). Three P1 clones, average size 80–100 kb, were identified: 9792, 9793 and 9794. PCR of 9792 using a series of primer pairs was conducted to confirm the sequence from oligonucleotide 10 070 to the site of polyadenylation (Fig. 2).

Placenta, uterus, and liver 'marathon ready' cDNA templates were used for 5' RACE (rapid amplification of cDNA ends) to obtain upstream sequence information. Twenty picomoles primer 10 070 and AP1, supplied with the 5' RACE amplification kit (Clontech), were used and the 5' RACE was carried out according to the manufacturer's instructions. The reactions were initially incubated at 94 °C for 1 min, followed by 35 cycles (94 °C, 20 s; 68 °C, 5 min) followed by a 10 min extension at 72 °C. The 5' RACE reaction resulted in a band of about 1.4 kb in all tissues.

### Northern blots

Northern blot analysis was performed using human brain, human fetal tissues and human multiple tissue Northern blots of poly A RNA from Clontech. The 415 bp DNA fragment was electrophoresed on a 1% agarose gel, the fragment was electroeluted, and then radioactively labeled using a random priming MEGAPRIME DNA labeling system (Amersham, Arlington Heights, IL, USA) according to the manufacturer's specifications. The probe was purified using NUCTRAP push column (Stratagene Cloning Systems, La Jolla, CA, USA). EXPRESSHYB (Clontech) solution was used for prehybridization and as a hybridizing solution for the Northern blots. Hybridization took place overnight at 68 °C, and the blots were then washed in 2 × SSC and 0.05% SDS at room temperature, followed by a wash in 0.1 × SSC and 0.1% SDS at 50 °C. The blots were also probed with an antisense riboprobe. The 415 bp PCR product created using primers 10 063 and 10 071 was subcloned into the p-GEM-T vector (Promega, Madison, WI, USA) according to the manufacturer's specifications. Orientation was determined to be in the sense direction under the Sp6 promoter. Riboprobe was synthesized using Sp6 RNA polymerase from the above plasmid. Human multiple tissue Northern blots I and II (Clontech) were prehybridized for 3 h at 55 °C. Radiolabeled Sp6 riboprobe was added to fresh hybridization solution at 3 × 10<sup>6</sup> c.p.m./ml and incubated for 36 h at 55 °C. The blots were then washed in 0.1 × SSC and 0.1% SDS at 60 °C and exposed to film.

## Results and Discussion

### EST homologs

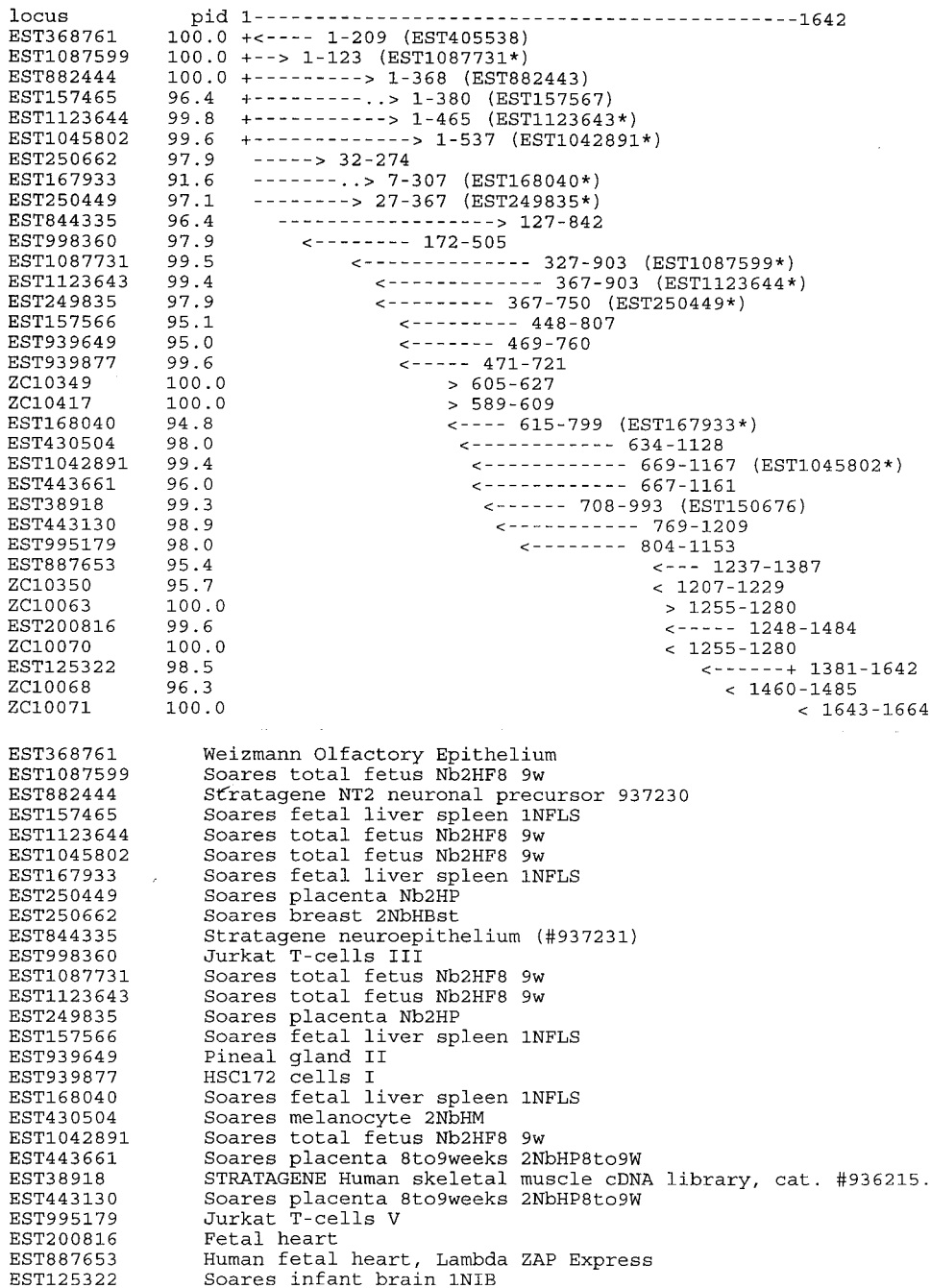
One EST (EST200 816) was identified as having acceptable homology to the putative Type II reptile EC3 (Troskie *et al.* 1998) and also to the region of TM6 to the end of TM7 of the Type I GnRH receptor. Another EST (EST157 566) was found with homology to the EC2 and TM5 regions. Surprisingly, although they are both annotated in dbEST as being 5'ESTs, they matched the Type I receptor in a reverse orientation (Figs 1 and 2). To explore this further, a longer virtual cDNA was created using the Rex program (see Materials and Methods) and EST157 566 as an anchor sequence. This revealed several more contiguous ESTs which contained coding and non-coding sequences, corresponding to regions of exons 2 and 3, and introns 1 and 2 of the Type I GnRH receptor gene (Kakar 1997) (Figs 1 and 2). All of these contiguous ESTs matched the assembled sequence in a frame opposite to their annotated direction, i.e. 5'ESTs matched in the reverse orientation, 3'ESTs matched in the forward orientation (see Fig. 1). In addition to the alignment of EST contigs shown in Fig. 1 we have found that the assembly can be extended, using EST data, by about 600 bases in

the 5' direction and about 200 bases in the 3' direction. These sequences appear to be intronic and 3' untranslated sequences respectively. These extending ESTs also matched the antisense transcript in a direction consistent with their clone annotation (data not shown). This supports the observation that all contiguous ESTs in the EST databases were derived from the antisense transcript.

### Gene structure, cDNA sequence and amino acid sequence

The EST sequences were used to design a series of oligonucleotides spanning the sequence extracted from the database and used for PCR amplification of cDNA prepared from mRNA from a range of human tissues. Selected PCR products, ESTs and the genomic clone (9792) were sequenced in both directions to confirm the entire 1642 bp sequence. The consensus sequence from the ESTs and our own sequencing (Fig. 2) reveals a nucleotide sequence which would encode a protein with homology to exon 2 (EC2 to intracellular loop domain 3 (IC3)) and exon 3 (IC3 to end of TM7) of the Type I GnRH receptor (Fig. 2). The overall amino acid identity in this region was 42% with greatest homology in the carboxyl terminal halves of EC2, IC3, TM6 and TM7, and the amino terminal half of TM5. This homology was much higher than with all other G-protein-coupled receptors (GPCRs) and similar to that of non-mammalian GnRH receptors (Troskie *et al.* 1998). Thus the homology with the GPCRs most closely related to the GnRH receptors (vasopressin, oxytocin and somatostatin 4 receptors) was 22–27% compared with 42–46% for GnRH receptors. The homology of EC3 of the human Type II GnRH receptor homolog to the EC3 of the lizard and *Xenopus* Type II GnRH receptors was 80% and 73% respectively (Troskie *et al.* 1998).

An interesting feature of the Type II GnRH receptor homolog is the presence of a carboxyl terminal tail which is absent in all of the mammalian Type I GnRH receptors (Eidne *et al.* 1992, Kaiser *et al.* 1992, Kakar *et al.* 1992, Reinhart *et al.* 1992, Tsutsumi *et al.* 1992, Brooks *et al.* 1993, Chi *et al.* 1993, Illing *et al.* 1993, Kakar *et al.* 1993, Perrin *et al.* 1993) but which is a universal feature in all other GPCRs and in non-mammalian GnRH receptors (Millar *et al.* 1997, Troskie *et al.* 1997a). The carboxyl terminal tail has been shown to play a role in homologous and heterologous desensitization of many GPCRs and its absence in the Type I GnRH receptor may be associated with a lack of rapid desensitization of ligand-stimulated inositol phosphate production (Davidson *et al.* 1994, McArdle *et al.* 1995, Heding *et al.* 1998). The presence of the carboxyl terminal tail in the Type II GnRH receptor homolog probably reflects the more ancient form of the receptor. This is supported by the recent demonstration of the presence of carboxyl terminal tails in GnRH receptors cloned from catfish (Tensen *et al.* 1997) and *Xenopus*,



**Figure 1** A contig map showing the location of different expressed sequence tags (EST prefix) and oligonucleotides (ZC prefix) on an anchor sequence which is the reverse complement of the antisense transcript. pid denotes the percentage identity of the EST to the anchor sequence; + refers to ESTs which have right or left overhang on the contig map. Arrowheads on the lines refer to the orientation of the EST with respect to the anchor. The numbers denote the range of match on the anchor sequence. In brackets are the names of the clone partners, (where available), derived from the dbEST annotations. An asterisk beside the EST name means that its clone partner is contained within the contig map. The tissues and cell line sources of the ESTs are indicated below.

```

1      GCATTGGAAAATGTTAAATTTTATTTTAAACATTCCTTACCATTATAGTFACTGCACATAAGACTAT
      TACTACTAAAGGTCACCTCAGAGTCCCTGCAAAAATGGCCCTGGAATTTGGCAGCACCCATTTACACAA
      TATTTCTTTTCCACAAAATAACAGACATACCAGGAAAATCATTTTCAGCTAAAATATGAGTGAGGTGG
      TAGAAATATCATCCCTTATAAAGCGCAATGTTAGAATAGTACTTGAGAAAAGCAGGATGTTTAAAGTTC
      CAGAATTAACAAACTTACTGTTTCAGCATCATATCAAGCCATAAAGGAAAGATAGGATTTTCAAGATAT
      ATTTCCAACCTTTTAAACATGGCACCATGGATGAAGTGTTCACAGCACTGTGCTGCTTCACTTGGAAAT
      TAAGATGAATTTGGGAGGACAGTATGACATAGTGGGTATGTTGGGTGGTGGGGGAACCGATCTTA
484    ATAGTCCTCAACTCCACTCCAGCTGTTCCTGTTCCACACGGTCCACTGAGCTGGCCAGTCCCTTTCAC
II     . A G P V P F T
I      L I Q L Y I F R M I H L A D S S G Q T K V F S
553    TCAGTGTGTACCAAAGGCAGCTTCAAGGCTCAATGGCAAGAGACCACCTATAACCTCTTACCTTCTG
II     Q C V T K G S F K A Q W Q E T Y N L F T F C
I      Q C V T H C S F S Q W W H Q A F Y N F F T F S
622    CTGCCCTCTTCTGCTGCCACTGACTGCCATGGCCATCTGCTATAGCCGCAITGCTCAGTGTGTCAG
II     C L L L L P L T A M A I C Y S R I V L S V S R
I      C L F I I P L F I M L I C N A K I I F T L T R
691    GCCCAGACAAGGAAGGGGAGCCATGGTGAGACTCCAATTCACAGCCCTTAATCCTTAACCCTAGTCCT
II     P Q T R K G S H G E T P I P R P .
I      V L H Q D P H
760    GTTGCCTTAGCATCATTTATTTATCTACCTACCTAATAGTATCTACCAGTCACTAAACCATGGTGAG
      ATCTAACCATGTCTAGCACCTGATGCTAGAGATAAATTTGGTGAATCCCTTCAATTTATAAACAGCTGA
      GTTAGCTGGACAAGGACTAGGGAGGCAATCAGTATATTATTTCTTGAACACCATCAAGTCTAGACTTG
      GTGGCTTCATATTTCTATATAAACCCCTGGGGGTAGAAATCATATAGTCCAGGTTGGGAAGGGGAAA
967    . S Q V G K G K
1036   ACGGTTTGCAACATTCCTCTCTTGTAGGAGGCGAGCTCTGTCTCACTAGCTATGCCCCCTCCATCAATT
II     T V C N I L L L V G G E L C L T S Y A P P S I
1105   CACCCATACTCAGATCAGAAGCTGAGTGTCTGAATTACAGTATATTTTCTAAATTCCTAGCCCTGCT
II     H P I L R S E A E C L N Y S I F S K F L A P A
1174   GGTGAATTTGCCCTCCCCCGCTCCTTTGACAATTTGCCCGGTGTTCTCGTCTCCGGCCCTGAGACTGGCC
II     G E F A L P R S F D N C P R V R L R A L R L A
I      E L Q L N Q S K N N I P R A R L K T L K M T
1243   CTGCTTAACTTACTGACCTTCATCCTCTGCTGGACACCTTATTACCTACTGGGTATGTTGGTACTGGTTC
II     L L N L L T F I L C W T P Y Y L L G M W Y W F
I      V A F A T S F T V C W T P Y Y V L G I W Y W F
1312   TCCCCCACCATGCTAACTGAAGTCCCTCCAGCCCTGAGCCACATCCTTTTCCTTGGGCCCTCCCAAT
II     S P T M L T E V P P S L S H I L F L L G M W Y W F
I      D P E M L N R L S D P V N H F F F L F A F L N
1381   GCTCCTTTGGATCCTCTCCTCTATGGGGCTTCACCCCTTGGCTGCCGAAGGGGACCCAGAACCTTAGT
II     A P L D P L L Y G A F T L G C R R G H Q E L S
I      P C F D P L I Y G Y F S L .
1450   ATAGACTTCTTAAAGAAGGGTCTGGGAGAATGCTCCAAGAGGAGATTTCATGCCTTTAGACAGCTGGAA
II     I D S S K E G S G R M L Q E E I H A F R Q L E
1519   GTACAAAAAAGCTGACATCAAGAAGGGCAGGAGAAAACAAAGGCATTTCTATAACATCTATCTGATCC
II     V Q K T V T S R R A G E T K G I S I T S I .
1588   TAACAGAGTATGTAGGAACAGAATAGTAAGTCTTTAGTGCATAAGATCTTAACA

```

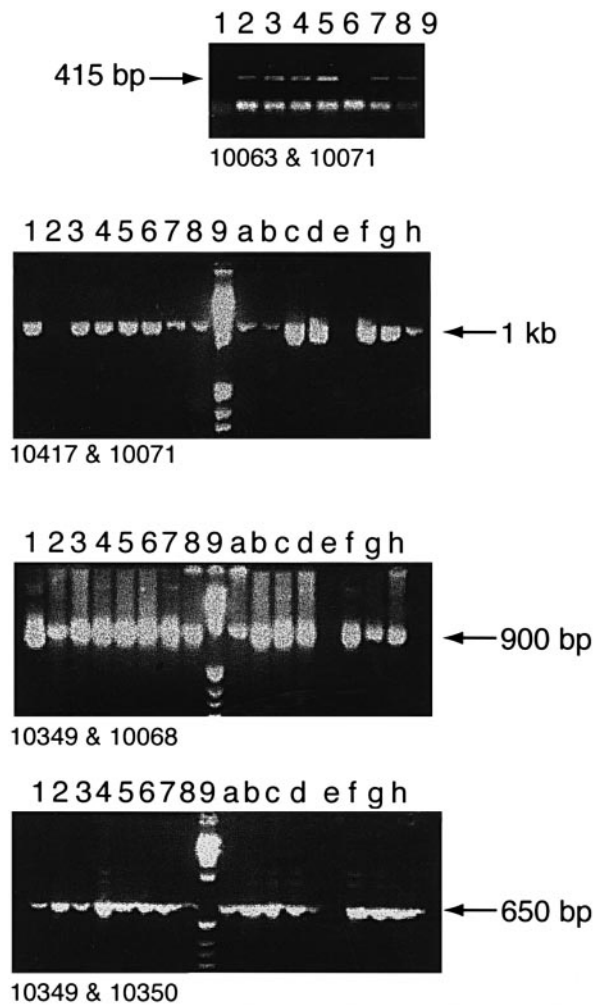


**Figure 2** Nucleotide sequences of the antisense transcript represented in the 3' to 5' orientation to reveal the encoded amino acid sequence of the putative Type II GnRH receptor and its similarity to the human Type I GnRH receptor. The alignment demonstrates the substantial amino acid identity in the presumed extracellular loop domains 2 and 3, intracellular loop 3 and transmembrane domains 5, 6 and 7 (underlined). In addition a COOH-terminal tail (C) is present in the Type II GnRH receptor. The polyadenylation consensus sequence (PA) and polyadenylation site (TTTT) are also indicated. The untranslated sequences correspond approximately to the positions of introns 1 and 2 of the Type I receptor genes. Putative exon 2 acceptor sites, the predicted intron 2 donor site, and putative exon 3 acceptor sites which would maintain the correct frame are shown in bold. Below are diagrammatic presentations of the structure of the human Type I GnRH receptor gene (Kakar 1997) and the putative Type II GnRH receptor gene. The positions of the extracellular (EC) and intracellular (IC) loop domains, transmembrane (TM) domains (dark blocks) and carboxyl terminal tail (C) are indicated.

chicken and goldfish (Millar *et al.* 1997, Troskie *et al.* 1997a, Illing *et al.* 1999).

The Type II GnRH receptor homolog does, however, have the DP motif in TM7 which is important functionally in the Type I GnRH receptor (Zhou *et al.* 1994) and which is more commonly NP in GPCRs (Millar *et al.* 1997, Sealson *et al.* 1997). It also has the identical motif QCVT containing the cysteine residue in EC2 which forms a critical disulfide bridge in the Type I receptor (Cook & Eidne 1997, Davidson *et al.* 1997). The sequence PRARL in the carboxyl terminal portion of IC3 is also retained but the alanine which is important for coupling (Myburgh *et al.* 1998) has been substituted by valine. Substitution of alanine with bulky and charged amino acids results in uncoupling of the Type I human GnRH receptor but a valine residue is acceptable and substitution only partially uncouples inositol phosphate production in the Type I receptor (Myburgh *et al.* 1998).

Between the putative exons 2 and 3 of the Type II GnRH receptor homolog is ~430 bp of sequence which corresponds in position to intron 2 of the Type I GnRH receptor (Kakar 1997) (Fig. 2). When read in the orientation encoding the putative Type II GnRH receptor (Fig. 2), there is only one possible donor site (the GT at position 717). There are four possible acceptor sites which retain the correct reading frame as indicated in Fig. 2. The last of these at position 1164 is closest to the analogous acceptor site of the Type I receptor gene and is a good candidate for processing if a transcript is indeed produced. Upstream of the putative exon 2 in the putative Type II GnRH receptor, the sequence is not homologous to exon 1 of the Type I receptor and encodes a stop (Fig. 2). It appears to be a non-coding sequence which corresponds approximately to the position of intron 1 of the Type I GnRH receptor (Kakar 1997). The 5' sequence of most of the PCR-generated cDNA comprised a poly T sequence. This, therefore, appeared to correspond to the poly A tail of the antisense transcript. This is confirmed by the presence of a polyadenylation signal sequence upstream of the poly A tail in the antisense transcript (Fig. 2). In some of the antisense transcripts a poly A tail is not incorporated in this position and a further 600 base extension is present which is not open reading frame in either orientation. This, therefore, appears to correspond to intron 1 of the Type I GnRH receptor (Kakar 1997), and to the 3' untranslated end of the antisense transcript (see later). Thus the putative intron 1 of the Type II GnRH receptor contains a polyadenylation signal for the antisense transcript. Sequencing of the genomic clones confirmed the polyadenylation signal sequence, an absence of the poly A and the 600 base extension. Extensive 5' RACE studies using a series of oligonucleotides complementary to the putative exon 2 of the Type II GnRH receptor and cDNA from a wide range of tissues failed to reveal any sequence homologous to exon 1 of the GnRH Type I receptor.



**Figure 3** PCR products of cDNA from various human tissues using oligonucleotide primers specific to the human putative Type II GnRH receptor. PCR was performed using the pairs of primers indicated and whose sequence is indicated from the position in Fig. 1. The human tissue cDNAs in the upper gel were: (1) lymph node, (2) fetal brain, (3) placenta, (4) uterus, (5) liver, (6) kidney, (7) spleen, (8) heart, (9) molecular weight markers and in the lower three gels: (1) ovary, (2) lymph node, (3) heart, (4) retina, (5) uterus, (6) placenta, (7) fetal brain I, (8) adult brain I, (9) molecular weight markers, (a) thyroid, (b) MCF 7 breast carcinoma cell line, (c) spinal cord, (d) fetal brain II, (e) blank, (f) hypothalamus, (g) adult brain II, (h) skeletal muscle. Similar bands were obtained with cerebellum, testis, insulinoma and pancreas in other experiments.

#### PCR and Northern blot analyses

cDNA from a range of human tissues and cell lines was subjected to PCR amplification using sets of oligonucleotide primers designed from the EST sequences. The majority of tissues gave intense single bands of the correct size (Fig. 3) whose identity was confirmed by sequencing and nested PCR or by Southern blotting and hybridization

with a labeled internal oligonucleotide (data not shown). This expression in many tissues accords with the wide range of tissues and cell lines in which the ESTs were found (Fig. 1). Since only antisense EST transcripts were found, donor and acceptor sites for sense RNA processing would not be recognized and intron sequences would fail to be spliced out. We confirmed this by demonstrating that PCR on various tissue cDNAs using primer pairs from putative exons 2 and exon 3 resulted in the production of the expected size products containing an additional 0.4 kb corresponding to intron 2 of the Type 1 GnRH receptor. These PCR products were not due to genomic DNA contamination as shown by control PCR reactions on 'cDNA' synthesized in the absence of reverse transcriptase as well as on DNase I-treated cDNA (data not shown).

Northern blot analysis confirmed these findings (Fig. 4). A strong signal was observed in all human tissues when the <sup>32</sup>P-labeled double-stranded 415 bp probe (see Materials and Methods) was used. Three transcript sizes were detected. Two transcript sizes were observed in all areas of the brain tested, at approximately 3 kb and 6 kb (Fig. 4). Signal intensity was highest for the 6 kb transcript. A third transcript size of 7.5 kb was present at varying intensity, with highest expression in cerebral cortex, medulla, spinal cord and corpus callosum. On the fetal blot, the highest level of expression of all three transcripts was in fetal kidney. In the multiple tissue blots expression of the 3 and 6 kb transcripts was highest in skeletal muscle, ovary, thyroid and spinal cord. The 7.5 kb transcript was more intense in spinal cord, brain, ovary, skeletal muscle and testis. A more stringent wash at 65 °C did not alter the banding patterns described above. However, when a single-stranded labeled riboprobe was used to detect a transcript for the Type II GnRH receptor no clear signal was detectable. The strong and universal signal obtained with the double stranded 415 bp probe therefore probably represents the antisense transcript. Its abundance and widespread occurrence in tissues is in keeping with encoding a ribonucleoprotein (see below).

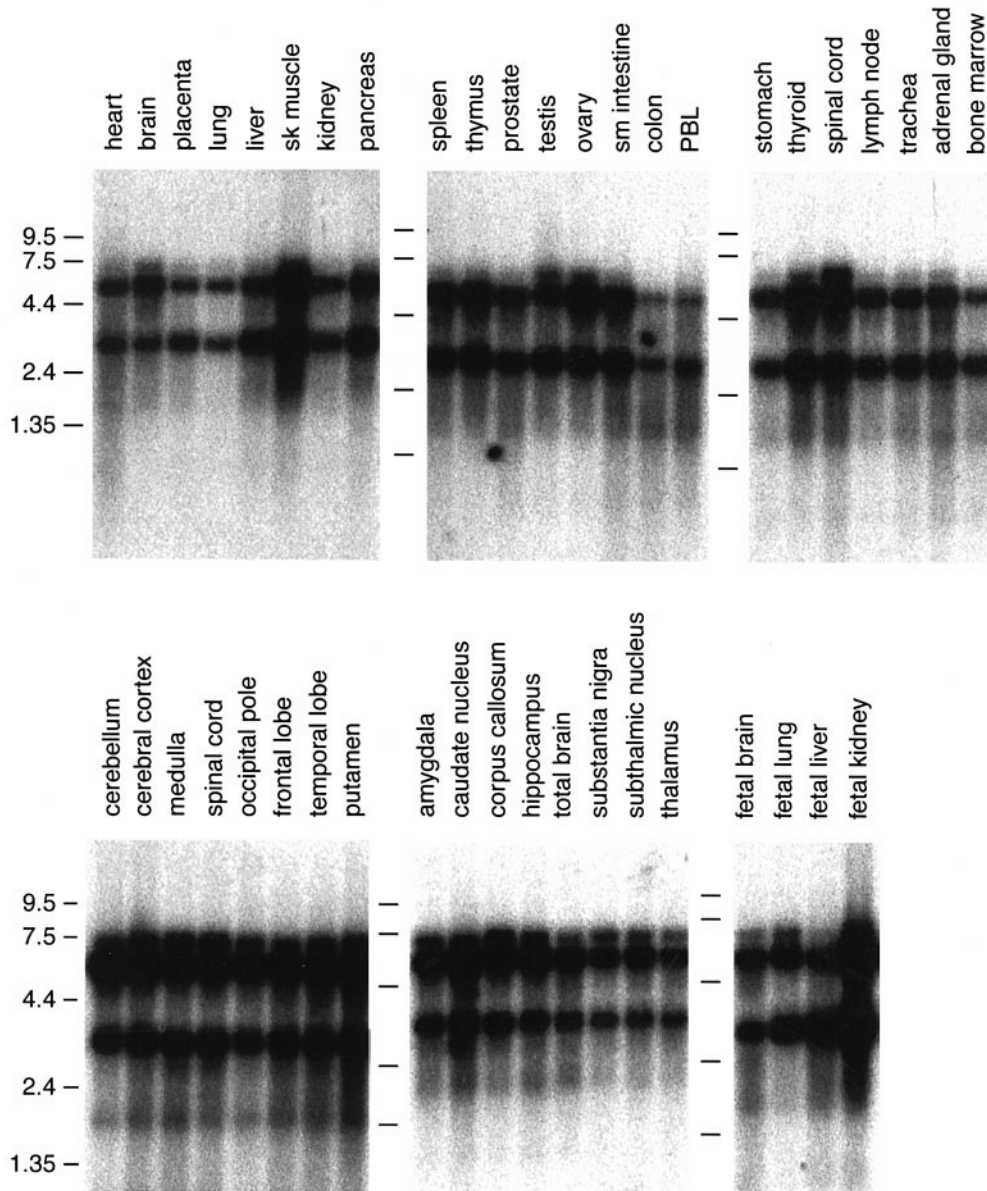
These findings demonstrate that sequence encoding a Type II GnRH receptor homolog of the pituitary Type I GnRH receptor exists in the human genome. The gene is highly expressed in a wide range of neural and non-neural adult and fetal tissues but the principle product is a transcript of the antisense DNA strand. The transcripts contained the equivalent of intron 2 of the GnRH Type I receptor (Kakar 1997) as potential donor and acceptor sites for splicing are only present in the sense orientation. These data also show that as transcription proceeds in the antisense direction into the equivalent of intron 1 of the Type I GnRH receptor, a polyadenylation signal is encountered which results in polyadenylation and the failure to transcribe the equivalent of the Type 1 GnRH receptor exon 1 if in fact it is present in the gene.

### Significance of the antisense transcript

Since the antisense transcripts are produced in abundance in all tissues examined, it was of interest to determine if they encoded a protein. However, none of the sequence revealed any significant open-reading frame and suggested that a functional protein is not translated. As this is unusual for a widely and abundantly expressed transcript, we undertook further exploration of the antisense transcript. Using the map in Fig. 1 we discovered that there are two 3'ESTs, EST882 444 and EST157 465, the 5' ends of which are not contained within this contig map or in the 3'-extended contig map. A Rex analysis of EST882 443 (the annotated 5' clone partner to EST882 444) revealed a transcript with homology to a *C. elegans* ribonucleoprotein. The Rex analysis of EST157 567 (the annotated 5' clone partner to EST157 465) produced a long virtual transcript of 3.5 kb. However, we have not explored this connection further, as based on annotated clone insert sizes, we believe that EST157 567 is incorrectly annotated as the clone partner to EST157 465, and that the correct clone partner to EST157 465 is EST157 566 which is within the contig map of our consensus sequence.

Our failure to detect a processed sense transcript coding for a full-length Type II GnRH receptor in a wide range of adult and fetal tissues and the absence of such a sense transcript in the EST database raises the possibility that the Type II GnRH receptor which we have demonstrated by binding studies in a non-mammal (Troskie *et al.* 1997b) is no longer expressed in eutherian mammals. If a functional Type II GnRH receptor gene is expressed, the very high expression of the antisense transcript is likely to inhibit translation due to RNA hybrid formation, and this feature would require mechanisms to counteract this. Alternatively, it is possible that our findings are explained by an extremely low abundance, a high degree of instability, or transient expression during limited stages of development. Another possibility is that this human putative Type II GnRH receptor may have been duplicated in the genome and we have been studying the duplicated version inserted behind the highly expressed ribonucleoprotein gene. However, chromosomal studies did not reveal a dual localization but only a single banding to 1q12-q21 (data not shown).

Transcription from the opposite DNA strand of a functional gene is a frequent feature in prokaryotes and can encode another functional protein or serve a regulatory role (Simons 1993). This phenomenon was initially thought not to occur in eukaryotes, but has recently been demonstrated in a number of genes (Knee & Murphy 1997). Antisense transcription has been implicated in regulating eukaryotic gene expression by inhibiting N-myc RNA elongation (Krystal *et al.* 1988), RNA editing for the glutamate receptor (Higuchi *et al.* 1993), mRNA splicing for the thyroid hormone receptor



**Figure 4** Northern blots of mRNA from human tissues. Blots were hybridized with the  $^{32}\text{P}$ -labeled 415 bp double stranded probe. Major 6 kb and 3 kb bands were present in all tissues while an additional 7.5 kb band was visible in cerebral cortex, medulla, spinal cord, corpus callosum, ovary, skeletal (sk) muscle, thyroid and testis. No clear signal was obtained on these blots when using a riboprobe to detect sense transcript. Small (sm) intestine; PBL, peripheral blood lymphocytes.

(Munroe & Lazar 1991), mRNA transport in the myelin deficient mouse gene (Tosic *et al.* 1990), mRNA half life for basic fibroblast growth factor (Kinelman & Kirschner 1987), and mRNA translation for the myosin heavy chain (Heywood 1986). The complementary pattern of expression of antisense and sense transcripts for the HoxaII (Hsieh-Li 1995) and basic fibroblast growth factor (Li *et al.* 1996) also suggest a regulatory role.

In other examples such as *Drosophila melanogaster*, cuticle protein (Henikoff *et al.* 1986) and decarboxylase (Spencer *et al.* 1986) genes, mouse tissue culture antisense transcripts (Williams & Fried 1986), cytochrome oxidase subunit III of *Trypanosoma brucei* (Volloch *et al.* 1991), the mouse globin gene (Volloch *et al.* 1996) and insulin-like growth factor-II (Sussenbach 1989), the role of antisense RNA transcription is unknown.



For some eukaryotic genes the antisense transcript encodes an open reading frame as in prokaryotes. *Xenopus* basic fibroblast growth factor (Li *et al.* 1996), *c-erbA* (Lazar *et al.* 1989, Miyajima *et al.* 1989) and GnRH (Adelman *et al.* 1987) fall into this category. Interestingly, the open reading frame antisense transcript of the GnRH gene is expressed in the heart (Adelman *et al.* 1987) but a translated protein has not been described. It has been suggested that the localization of GnRH antisense RNA in the nucleus may indicate a role in GnRH mRNA processing (Knee & Murphy 1997).

The present demonstration of the transcription of the opposite strand to a putative novel GnRH receptor homolog gene adds to the expanding series of examples of antisense transcriptions. Although a full length sense transcript which would encode a Type II GnRH receptor in human tissue was not found, the presence of this sequence in the human genome demonstrates that at least this receptor existed at some stage of vertebrate evolution. Binding studies demonstrating a second type of GnRH receptor with distinctly different pharmacology in *Xenopus* sympathetic ganglion membranes (Troskie *et al.* 1997b) supports this concept. We have sequenced a partial clone of a GnRH receptor in the *Xenopus* sympathetic ganglion whose sequence differs from the *Xenopus* pituitary receptor (B E Troskie, T R Ott, C A Flanagan, J P Hapgood, R P Millar and N Illing, unpublished data), and which is identical to the genomic EC3 sequence of the designated Type II GnRH receptor (Troskie *et al.* 1998). This clone has 54% overall amino acid identity with the human sequence described here and 73% identity in the ligand-selective EC3 domain. The cloning of the *Xenopus* and other non-mammalian receptors will assist in determining the phylogenetic origins, expression and functions of this intriguing putative Type II GnRH receptor. It should also provide the key to a Type II GnRH receptor which might exist in mammals or establish that the current gene we have identified is a vestigial version of the non-mammalian Type II receptor.

## References

- Adelman JP, Bond CT, Douglass J & Herbert E 1987 Two mammalian genes transcribed from opposite strands of the same DNA locus. *Science* **235** 1514–1517.
- Altschul SF, Warren G, Miller W, Myers EW & Lipman DJ 1990 Basic local alignment search tool. *Journal of Molecular Biology* **215** 403–410.
- Brooks J, Taylor PL, Saunders PT, Eidne KA, Struthers WJ & McNeilly AS 1993 Cloning and sequencing of the sheep pituitary gonadotropin-releasing hormone receptor and changes in expression of its mRNA during the estrous cycle. *Molecular and Cellular Endocrinology* **94** R23–R27.
- Chi L, Zhou W, Prikhozhan A, Flanagan C, Davidson JS, Golembo M, Illing N, Millar RP & Sealfon SC 1993 Cloning and characterization of the human GnRH receptor. *Molecular and Cellular Endocrinology* **91** R1–R6.
- Cook JV & Eidne K 1997 An intramolecular disulfide bond between conserved extracellular cysteines in the gonadotropin-releasing hormone receptor is essential for binding and activation. *Endocrinology* **138** 2800–2806.
- Davidson JS, Wakefield IK & Millar RP 1994 Absence of rapid desensitization of the mouse gonadotropin-releasing hormone receptor. *Biochemical Journal* **300** 299–302.
- Davidson JS, Assefa D, Pawson A, Davies P, Hapgood J, Becker I, Flanagan C, Roeske R & Millar R 1997 Irreversible activation of the gonadotropin-releasing hormone receptor by photoaffinity cross-linking: localization of attachment site to Cys residue in the N-terminal segment. *Biochemistry* **36** 12881–12889.
- Eidne KA, Sellar RE, Couper G, Anderson L & Taylor PL 1992 Molecular cloning and characterisation of the rat pituitary gonadotropin-releasing hormone (GnRH) receptor. *Molecular and Cellular Endocrinology* **90** R5–R9.
- Flanagan CA, Becker II, Davidson JS, Wakefield IK, Zhou W, Sealfon SC & Millar RP 1994 Glutamate<sup>301</sup> of the mouse gonadotropin-releasing hormone receptor confers specificity for arginine<sup>8</sup> of mammalian gonadotropin-releasing hormone. *Journal of Biological Chemistry* **269** 22636–22641.
- Heding A, Vrecl M, Bogerd J, McGregor A, Sellar R, Taylor PL & Eidne KA 1998 Gonadotropin releasing hormone receptors with intracellular carboxyl terminal tails undergo acute desensitization of total inositol phosphate production and exhibit accelerated internalisation kinetics. *Journal of Biological Chemistry* **273** 11472–11477.
- Henikoff S, Keene MA, Fechtel K & Fristron JW 1986 Gene within a gene: nested *Drosophila* genes encode unrelated protein on opposite DNA strands. *Cell* **44** 33–42.
- Heywood SM 1986 tcRNA as a naturally occurring antisense RNA in eukaryotes. *Nucleic Acids Research* **14** 6771–6772.
- Higuchi M, Single FN, Kohler M, Sommer B, Sprengel R & Seeburg PH 1993 RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. *Cell* **75** 1361–1370.
- Hsieh-Li HM, Witte DP, Weinstein M, Bradford W, Li H, Small K & Potter SS 1995 *Hoxa11* structure, extensive antisense transcription and function in male and female fertility. *Development* **121** 1373–1385.
- Illing N, Jacobs GF, Becker II, Flanagan CA, Davidson JS, Eales A, Zhou W, Sealfon SC & Millar RP 1993 Comparative sequence analysis and functional characterization of the cloned sheep gonadotropin-releasing hormone receptor reveal differences in primary structure and ligand specificity among mammalian receptors. *Biochemical and Biophysical Research Communications* **196** 745–751.
- Illing N, Troskie B, Nahorniak C, Hapgood J, Peter RE & Millar RP 1999 Two gonadotropin-releasing hormone receptor subtypes with distinct ligand selectivity and differential distribution in brain and pituitary in the goldfish (*Carassius auratus*). *Proceedings of the National Academy of Sciences of the USA* **96** 2526–2531.
- Kaiser UB, Zhao D, Cardona GR & Chin WW 1992 Isolation and characterization of cDNAs encoding the rat pituitary gonadotropin-releasing hormone receptor. *Biochemical and Biophysical Research Communications* **189** 1645–1652.
- Kakar SS 1997 Molecular structure of the human gonadotropin-releasing hormone receptor gene. *European Journal of Endocrinology* **137** 183–192.
- Kakar SS, Musgrove LC, Devor DC, Sellers JC & Neill JD 1992 Cloning, sequencing, and expression of human gonadotropin releasing hormone (GnRH) receptor. *Biochemical and Biophysical Research Communications* **189** 289–295.
- Kakar SS, Rahe CH & Neill JD 1993 Molecular cloning, sequencing, and characterizing the bovine receptor for gonadotropin releasing hormone (GnRH). *Domestic Animal Endocrinology* **10** 335–342.

- Kineman D & Kirschner M 1987 Synergistic induction of mesoderm by FGF and TGF- $\beta$  and the identification of a mRNA coding for FGF in the early *Xenopus* embryo. *Cell* **51** 869–877.
- King JA & Millar RP 1979 Heterogeneity of vertebrate luteinizing hormone-releasing hormone. *Science* **206** 67–69.
- King JA & Millar RP 1997 Coordinated evolution of GnRHs and their receptors. In *GnRH Neurons: Gene to Behavior*, pp 51–77. Eds IS Parhar and Y Sakuma. Tokyo: Brain Shuppan Publishers.
- Knee R & Murphy PR 1997 Regulation of gene expression by natural antisense RNA transcripts. *Neurochemistry International* **31** 379–392.
- Krystal G, Birrer M, Way J, Nau M, Sausville E, Thompson C, Minna J & Bettey J 1988 Multiple mechanisms of transcriptional regulation of the *myc* gene family in small cell lung cancer. *Molecular and Cellular Biology* **8** 3373–3381.
- Lazar MA, Hodin RA, Darling DS & Chin WW 1989 A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat *c-erbA* transcriptional unit. *Molecular and Cellular Biology* **9** 1128–1136.
- Li AW, Too CKL & Murphy PR 1996 The basic fibroblast growth factor (FGF-2) antisense RNA (GFG) is translated into MutT-related protein *in vivo*. *Biochemical and Biophysical Research Communications* **223** 19–23.
- McArdle CA, Forrest-Owen W, Willars G, Davidson J, Poch A & Kratzmeier M 1995 Desensitization of gonadotropin-releasing hormone action in the gonadotrope-derived alpha T3-1 cell line. *Endocrinology* **136** 4864–4871.
- Millar RP & King JA 1988 Evolution of gonadotropin-releasing hormone: multiple usage of a peptide. *News Physiological Science* **3** 49–53.
- Millar RP, Troskie B, Sun YM, Ott T, Wakefield I, Myburgh D, Pawson A, Davidson JS, Flanagan C, Katz A, Hapgood J, Illing N, Weinstein H, Sealfon SC, Peter RE, Terasawa E & King JA 1997 Plasticity in the structural and functional evolution of GnRH: a peptide for all seasons. In *Proceedings of the Thirteenth International Congress of Comparative Endocrinology* pp 15–27. Eds S Kawashima & S Kikuyama. Bologna: Monduzzi Editore.
- Munroe SH & Lazar MA 1991 Inhibition of *c-erbA* mRNA splicing by a naturally occurring antisense RNA. *Journal of Biological Chemistry* **266** 22083–22086.
- Miyajima N, Horiuchi R, Shibuya Y, Fukushige S, Matsubara K, Toyoshima K & Yamamoto T 1989 Two *erbA* homologs encoding proteins with different T3 binding capacities are transcribed from opposite DNA strands of the same genetic locus. *Cell* **57** 31–39.
- Myburgh DB, Millar RP & Hapgood JP 1998 Alanine 261 in intracellular loop 3 of the human gonadotropin-releasing hormone receptor is crucial for G-protein coupling and receptor internalisation. *Biochemical Journal* **331** 893–896.
- Pearson W, Wood T, Zhang Z & Miller W 1997 Comparison of DNA sequences with protein sequences. *Genomics* **46** 24–36.
- Perrin MH, Bilezikjian LM, Hoeger C, Donaldson CJ, Rivier J, Haas Y & Vale WW 1993 Molecular and functional characterization of GnRH receptors cloned from rat pituitary and a mouse pituitary tumor cell line. *Biochemical and Biophysical Research Communications* **191** 1139–1144.
- Reinhart J, Mertz LM & Catt KJ 1992 Molecular cloning and expression of cDNA encoding the murine gonadotropin-releasing hormone receptor. *Journal of Biological Chemistry* **267** 21281–21284.
- Sealfon SC, Weinstein H & Millar RP 1997 Molecular mechanisms of ligand interaction with the gonadotropin-releasing hormone receptor. *Endocrine Reviews* **18** 180–205.
- Sherwood NM, Lovejoy DA & Coe IR 1993 Origin of mammalian gonadotropin-releasing hormones. *Endocrine Reviews* **14** 241–254.
- Simons RW 1993 The control of prokaryotic and eukaryotic gene expression by naturally occurring antisense RNA. In *Antisense Research and Applications* pp 97–124. Eds ST Crooke & M LeBlew. Boca Raton: CRC Press Inc.
- Spencer CA, Gietz RD & Hodgetts RB 1986 Overlapping transcription units dopa decarboxylase region of *Drosophila*. *Nature* **322** 279–281.
- Sussenbach JS 1989 The gene structure of the insulin-like growth factor family. *Progress in Growth Factor Research* **1** 33–48.
- Tensen C, Okuzawa K, Blomenrohr M, Rebers F, Leurs R, Bogerd J, Schultz R & Goos H 1997 Distinct efficacies of two endogenous ligands on a single cognate gonadoliberm receptor. *European Journal of Biochemistry* **243** 134–140.
- Tosic M, Roach A, de Rivaz JC, Dolivo M & Matthieu JM 1990 Post-transcriptional events are responsible for low expression of myelin basic protein in myelin deficient mice: role of natural antisense RNA. *EMBO Journal* **9** 401–406.
- Troskie BE, Sun Y-M, Hapgood JP, Sealfon SC, Illing N & Millar RP 1997a Mammalian GnRH receptor functional features revealed by comparative sequences of goldfish, frog and chicken receptors. *79th Annual Meeting of the Endocrine Society*, June 11–14 1997, Minneapolis, Minnesota, Abstract P1–130.
- Troskie B, King JA, Millar RP, Peng Y-Y, Kim J, Figueras H & Illing N 1997b Chicken GnRH II-like peptides and a GnRH receptor selective for chicken GnRH II in amphibian sympathetic ganglia. *Neuroendocrinology* **65** 396–402.
- Troskie B, Sun Y-M, Rumbak E, Hapgood J, Swanson P, Sealfon SC, Conklin D & Millar R 1998 Identification of three putative GnRH receptor sub-types in vertebrates. *General and Comparative Endocrinology* **112** 296–302.
- Tsutsumi M, Zhou W, Millar RP, Mellon PL, Roberts JL, Flanagan CA, Dong K, Gillo B & Sealfon SC 1992 Cloning and functional expression of a mouse gonadotropin-releasing hormone receptor. *Molecular Endocrinology* **6** 1163–1169.
- Volloch V, Schweitzer B & Rits S 1991 Identification of negative-strand complements to cytochrome oxidase subunit III RNA in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences of the USA* **93** 10671–10675.
- Volloch V, Schweitzer B & Rits S 1996 Antisense globin RNA in mouse erythroid tissues: structure, origin, and possible function. *Proceedings of the National Academy of Sciences of the USA* **93** 2476–2481.
- Williams T & Fried M 1986 A mouse locus at which transcription from both DNA strands produces mRNAs complementary at their 3' ends. *Nature* **322** 275–279.
- Yee D & Conklin D 1998 Automated clustering and assembly of large EST collections. *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology*, pp 203–211. Eds J Glasgow, T Littlejohn, F Major, R Lathrop, D Sankoff & C Sensen. Montreal: AAAI Press.
- Zhou W, Flanagan C, Ballesteros JA, Konvicka K, Davidson JS, Weinstein H, Millar RP & Sealfon SC 1994 A reciprocal mutation supports helix 2 and helix 7 proximity in the gonadotropin-releasing hormone receptor. *Molecular Pharmacology* **45** 165–170.

Received 18 November 1998

Accepted 22 February 1999