Vol. 205, No. 3, 1994 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
December 30, 1994 Pages 1976-1983

RORγ: THE THIRD MEMBER OF ROR/RZR ORPHAN RECEPTOR SUBFAMILY THAT IS HIGHLY EXPRESSED IN SKELETAL MUSCLE[1]

Takahisa Hirose[2], Robert J. Smith, and Anton M. Jetten[3]

Cell Biology Section, Laboratory of Pulmonary Pathobiology, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709

Received November 17, 1994

In this study, we describe the identification and cloning of a novel member of the nuclear receptor superfamily. This orphan receptor, referred to as RORγ, belongs to the ROR/RZR subfamily. The open reading frame of RORγ encodes a protein of 560 amino acid residues with a predicted molecular mass of 63 kD. The amino acid sequence of RORγ exhibits a 50 and 51% identity with those of RORα/RZRα and RZRβ, respectively, whereas the DNA-binding domains were 89% identical. RORγ was localized on human chromosome 1. Northern blot analysis using RNA from multiple tissues indicated that RORγ is expressed in several tissues but is most highly expressed in skeletal muscle.

The nuclear receptor gene superfamily encodes an increasing number of transcriptional regulators that play critical roles during homeostasis and specific stages of development (1,2). Members of this gene family includes the steroid hormone, thyroid hormone and retinoid receptors, and orphan receptors for which a ligand has not yet been identified (3,4). The members of this family share a common modular structure that includes a highly homologous DNA-

[1]Sequence data from this article have been deposited with EMBL/GenBank Data Libraries under Accession Number U16997.

[2]Current address: Department of Medicine III, Osaka University Medical School, Yamadaoka, Suita, Osaka, Japan.

[3]To whom reprint requests should be addressed. Fax.: 919-541-4133; E-mail: jetten@niehs.nih.gov.

0006-291X/94 $5.00

binding domain (DBD) containing two "zinc-finger" motifs. Most of these nuclear receptors bind as monomers (5,6), homodimers and heterodimers to response elements composed of the single half-sitemotif PuGGTCA preceded by an AT-rich sequence or direct, palindromic or inverted palindromic repeats of the core motif spaced by one or more nucleotides (reviews in 7,8).

Using reverse transcription-PCR and degenerate primers whose designs were based on the two most conserved regions of the DBD of known members of the nuclear receptor superfamily, we identified and cloned a novel orphan receptor, named RORγ. Based on its overall amino acid homology (50-51%) with RORα/RZRα and RZRβ, this gene encodes the third member of the ROR/RZR subfamily of orphan receptors (6,9,10). Each member of the ROR/RZR subfamily exhibits a different tissue distribution suggesting that they have different functions. RORγ appears to play a role in controlling gene expression in several tissues and particularly in skeletal muscle.

MATERIALS AND METHODS

PCR amplification - A set of degenerate primers was designed according to the most highly conserved sequence of the DNA-binding domain of members of the nuclear receptor family as previously described (11,12). Single strand complementary DNA reverse transcribed from human pancreas poly(A)$^+$ RNA was employed with the primers in the amplification reaction using the Amplitaq kit and a DNA thermal cycler (Perkin-Elmer Cetus). The DNA products of 130bp were isolated from the gel and directly ligated to the TA cloning vector (Invitrogen). The DNA inserts of thirty clones were analyzed by double strand dideoxy DNA sequencing using Sequenase (U.S.Biochemical). One novel DBD sequence was obtained and named RORγ-DBD.

Anchor PCR - The 5'-RACE kit (BRL) was employed to isolate a greater region of the gene encoding RORγ (11,12). Briefly, two sequential antisense primers were designed from the sequence of the RORγ-DBD PCR fragment. Single strand cDNA was synthesized from 1μg of human pancreas poly(A)$^+$RNA with Superscript reverse transcriptase using the first strand primer according to the manufacturer's recommendation (BRL). Following denaturation by heating, the products were homopolymerically tailed with deoxy-CTP and terminal deoxytransferase and subsequently amplified by PCR at stringent conditions with the anchor primer poly-(dG) and the nested primer. The amplification products were subcloned into pBluescript plasmid (Stratagene) and 3 independent clones sequenced.

cDNA library screening - A λgt10-phage cDNA library from human skeletal muscle (Clontech) was screened under stringent conditions

with a multiprimed cDNA probe made from the anchor PCR products. cDNA inserts from several positive clones were subcloned into the *EcoR*I site of pBluescript. The complete DNA sequences of both strands were determined by the method described above.

Chromosomal localization - A BIOS Blot Somatic Cell Hybrid Panel was purchased from BIOS Laboratories (New Haven, CT). The blot was made from *Taq*I digested genomic DNA prepared from different human/hamster or human/mouse hybrid cell lines. Each cell line contained in addition to a complete set of hamster or mouse chromosomes one or more human chromosomes. The human chromosomes contained in each hybrid cell line are listed in Fig. 3B. The cell line 016 represents a human/mouse hybrid cell line, all others are human/hamster hybrid cell lines. The blot was hybridized with ^{32}P-labelled hRORγ probe (nt 1-250) and washed with 0.2x SSC and 0.1% SDS at 65°C.

Northern blot analysis - Human multiple tissue Northern blots were purchased from Clontech. The blots were hybridized with a ^{32}P-labeled (Amersham) RORγ probe encoding the nucleotides 1 to 250 as described previously (13).

RESULTS AND DISCUSSION

Cloning of RORγ cDNA - Highly degenerate primers were designed according to the two best conserved amino acid sequences present in the DBD of members of the nuclear receptor family. These primers were then employed in amplification reactions with single strand cDNA from human pancreas poly(A)$^+$ RNA in order to identify DBD's of novel nuclear receptors. The amplified fragments of the expected size (130 bp) were cloned in TA vector and 30 clones sequenced. Most of the cDNA sequences encoded DBD's of known receptors including those of RARα (14,15) and RXRβ (16). Among these cDNAs, we identified one which encoded a unique sequence not previously described. The amino acid residues predicted by this cDNA sequence suggested that it encoded the DBD of a novel member of the nuclear receptor superfamily which we refer to as RORγ. With the help of the 5'-RACE method we obtained a 400 bp fragment of the RORγ gene that was subsequently used as a probe in screening a human skeletal muscle cDNA library. After screening 9 x10^5 independent plaques, three positive clones were obtained which combined included almost the full length coding region of RORγ. Analysis of this RORγ sequence revealed a long open reading frame which starts with a putative initiation codon at nucleotide 70 and terminates with a putative stop codon at 1750 (Fig. 1). On this basis, RORγ encodes a protein of 560 amino acid residues with a predicted molecular weight of 62.6 kilodaltons. Fig. 2A shows the amino acid comparison of RORγ and

CCCCTGGGCCCTGCTCCCTGCCCTCCTGGGCAGCCAGGGCAGCCAGGACGGCACCAAGGGAGCTGCCCC 69

ATG GAC AGG GCC CCA CAG AGA CAG CAC CGA GCC TCA CGG GAG CTG CTG GCT GCA AAG AAG 129
Met Asp Arg Ala Pro Gln Arg Gln His Arg Ala Ser Arg Glu Leu Leu Ala Ala Lys Lys 20

ACC CAC ACC TCA CAA ATT GAA GTG ATC CCT TGC AAA ATC TGT GGG GAC AAG TCG TCT GGG 189
Thr His Thr Ser Gln Ile Glu Val Ile Pro Cys Lys Ile Cys Gly Asp Lys Ser Ser Gly 40

ATC CAC TAC GGG GTT ATC ACC TGT GAG GGG TGC AAG GGC TTC TTC CGC CGG AGC CAG CGC 249
Ile His Tyr Gly Val Ile Thr Cys Glu Gly Cys Lys Gly Phe Phe Arg Arg Ser Gln Arg 60

TGT AAC GCG GCC TAC TCC TGC ACC CGT CAG CAG AAC TGC CCC ATC GAC CGC ACC AGC CGA 309
Cys Asn Ala Ala Tyr Ser Cys Thr Arg Gln Gln Asn Cys Pro Ile Asp Arg Thr Ser Arg 80

AAC CGA TGC CAG CAC TGC CGC CTG CAG AAA TGC CTG GCG CTG GGG ATG TCC CGA GAT GCT 369
Asn Arg Cys Gln His Cys Arg Leu Gln Lys Cys Leu Ala Leu Gly Met Ser Arg Asp Ala 100

GTC AAG TTC GGC CGC ATG TCC AAG AAG CAG AGG GAC AGC CTG CAT GCA GAA GTG CAG AAA 429
Val Lys Phe Gly Arg Met Ser Lys Lys Gln Arg Asp Ser Leu His Ala Glu Val Gln Lys 120

CAG CTG CAG CAG CGG CAA CAG CAG CAA CAG GAA CCA GTG GTC AAG ACC CCT CCA GCA GGG 489
Gln Leu Gln Gln Arg Gln Gln Gln Gln Gln Glu Pro Val Val Lys Thr Pro Pro Ala Gly 140

GCC CAA GGA GCA GAT ACC CTC ACC TAC ACC TTG GGG CTC CCA GAC GGG CAG CTG CCC CTG 549
Ala Gln Gly Ala Asp Thr Leu Thr Tyr Thr Leu Gly Leu Pro Asp Gly Gln Leu Pro Leu 160

GGC TCC TCG CCT GAC CTG CCT GAG GCT TCT GCC TGT CCC CCT GGC CTC CTG AAA GCC TCA 609
Gly Ser Ser Pro Asp Leu Pro Glu Ala Ser Ala Cys Pro Pro Gly Leu Leu Lys Ala Ser 180

GGC TCT GGG CCC TCA TAT TCC AAC AAC TTG GCC AAG GCA GGG CTC AAT GGG GCC TCA TGC 669
Gly Ser Gly Pro Ser Tyr Ser Asn Asn Leu Ala Lys Ala Gly Leu Asn Gly Ala Ser Cys 200

CAC CTT GAA TAC AGC CCT GAG CGG GGC AAG GCT GAG GGC AGA GAG AGC TTC TAT AGC ACA 729
His Leu Glu Tyr Ser Pro Glu Arg Gly Lys Ala Glu Gly Arg Glu Ser Phe Tyr Ser Thr 220

GGC AGC CAG CTG ACC CCT GAC CGA TGT GGA CTT CGT TTT GAG GAA CAC AGG CAT CCT GGG 789
Gly Ser Gln Leu Thr Pro Asp Arg Cys Gly Leu Arg Phe Glu Glu His Arg His Pro Gly 240

CTT GGG GAA CTG GGA CAG GGC CCA GAC AGC TAC GGC AGC CCC AGT TTC CGC AGC ACA CCG 849
Leu Gly Glu Leu Gly Gln Gly Pro Asp Ser Tyr Gly Ser Pro Ser Phe Arg Ser Thr Pro 260

GAG GCA CCC TAT GCC TCC CTG ACA GAG ATA GAG CAC CTG GTG CAG AGC GTC TGC AAG TCC 909
Glu Ala Pro Tyr Ala Ser Leu Thr Glu Ile Glu His Leu Val Gln Ser Val Cys Lys Ser 280

TAC AGG GAG ACA TGC CAG CTG CGG CTG GAG GAC CTG CTG CGG CAG CGC TCC AAC ATC TTC 969
Tyr Arg Glu Thr Cys Gln Leu Arg Leu Glu Asp Leu Leu Arg Gln Arg Ser Asn Ile Phe 300

TCC CGG GAG GAA GTG ACT GGC TAC CAG AGG AAG TCC ATG TGG GAG ATG TGG GAA CGG TGT 1029
Ser Arg Glu Glu Val Thr Gly Tyr Gln Arg Lys Ser Met Trp Glu Met Trp Glu Arg Cys 320

GCC CAC CAC CTC ACC GAG GCC ATT CAG TAC GTG GTG GAG TTC GCC AAG AGG CTC TCA GGC 1089
Ala His His Leu Thr Glu Ala Ile Gln Tyr Val Val Glu Phe Ala Lys Arg Leu Ser Gly 340

TTT ATG GAG CTC TGC CAG AAT GAC CAG ATT GTG CTT CTC AAA GCA GGA GCA ATG GAA GTG 1149
Phe Met Glu Leu Cys Gln Asn Asp Gln Ile Val Leu Leu Lys Ala Gly Ala Met Glu Val 360

GTG CTG GTT AGG ATG TGC CGG GCC TAC AAT GCT GAC AAC CGC ACG GTC TTT TTT GAA GGC 1209
Val Leu Val Arg Met Cys Arg Ala Tyr Asn Ala Asp Asn Arg Thr Val Phe Phe Glu Gly 380

AAA TAC GGT GGC ATG GAG CTG TTC CGA GCC TTG GGC TGC AGC GAG CTC ATC AGC TCC ATC 1269
Lys Tyr Gly Gly Met Glu Leu Phe Arg Ala Leu Gly Cys Ser Glu Leu Ile Ser Ser Ile 400

TTT GAC TTC TCC CAC TCC CTA AGT GCC TTG CAC TTT TCC GAG GAT GAG ATT GCC CTC TAC 1329
Phe Asp Phe Ser His Ser Leu Ser Ala Leu His Phe Ser Glu Asp Glu Ile Ala Leu Tyr 420

ACA GCC CTT GTT CTC ATC AAT GCC CAT CGG CCA GGG CTC CAA GAG AAA AGG AAA GTA GAA 1389
Thr Ala Leu Val Leu Ile Asn Ala His Arg Pro Gly Leu Gln Glu Lys Arg Lys Val Glu 440

CAG CTG CAG TAC AAT CTG GAG CTG GCC TTT CAT CAT CAT CTC TGC AAG ACT CAT CGC CAA 1449
Gln Leu Gln Tyr Asn Leu Glu Leu Ala Phe His His His Leu Cys Lys Thr His Arg Gln 460

AGC ATC CTG GCA AAG CTG CCA CCC AAG GGG AAG CTT CGG AGC CTG TGT AGC CAG CAT GTG 1509
Ser Ile Leu Ala Lys Leu Pro Pro Lys Gly Lys Leu Arg Ser Leu Cys Ser Gln His Val 480

GAA AGG CTG CAG ATC TTC CAG CAC CTC CAC CCC ATC GTG GTC CAA GCC GCT TTC CCT CCA 1569
Glu Arg Leu Gln Ile Phe Gln His Leu His Pro Ile Val Val Gln Ala Ala Phe Pro Pro 500

CTC TAC AAG GAG CTC TTC AGC ACT GAA ACC GAG TCA CCT GTG GGC TGT CCA AGT GAC CTG 1629
Leu Tyr Lys Glu Leu Phe Ser Thr Glu Thr Glu Ser Pro Val Gly Cys Pro Ser Asp Leu 520

GAA GAG GGA CTC CTT GCC TCT CCC TAT GGC CTG CTG GCC ACC TCC CTG GAC CCC GTT CCA 1689
Glu Glu Gly Leu Leu Ala Ser Pro Tyr Gly Leu Leu Ala Thr Ser Leu Asp Pro Val Pro 540

CCC TCA CCC TTT TCC TTT CCC ATG AAC CCT GGA GGG TGG TCC CCA CCA GCT CTT TGG AAG 1749
Pro Ser Pro Phe Ser Phe Pro Met Asn Pro Gly Gly Trp Ser Pro Pro Ala Leu Trp Lys 560

TGA*GCAGATGCTGCGGCTGGCTTTCTGTCAGCAGGCCGGCCTGGCAGTGGGACAATCGCCAGAGGGTGGG 1819

Fig. 1. Nucleotide and deduced amino acid sequence of RORγ. Nucleotides and amino acids are numbered on the right side of the sequence. The putative initiation codon is at nt 70. The DNA-binding domain and putative ligand binding domain are underlined. The termination codon (nt 1750) at the end of RORγ amino acid sequence is indicated as *.

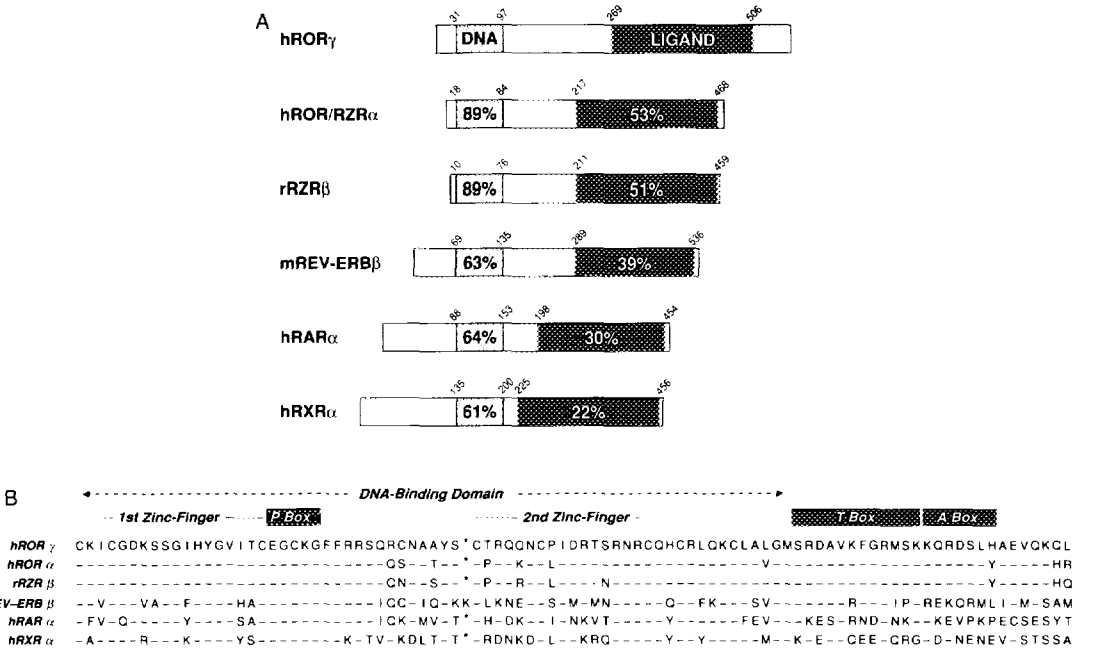

Fig. 2. Comparison between the amino acid sequence of hRORγ and several other members of the nuclear receptor family. A. Alignment of the amino acid sequences of the DBD's and ligand-binding domains. The numbers in each box indicate the percentage of identity with hRORγ. The DBD and ligand-binding domain are indicated. B. Comparison between the amino acid sequence of the DBD, P-, T- and A-box of RORγ, ROR/RZRα, RZRβ, Rev-erbβ, RARα and RXRα. Specific references are ROR/RZRα (6,9), RZRβ (10), Rev-erbβ (17), RARα (14,15) and RXRα (18). The prefix h, m, and r denotes human, mouse, and rat, respectively. * indicates gap.

several other nuclear receptors. The amino acid sequence of RORγ was most homologous (respectively, 51 and 50% identity) to that of hRZRα/RORα and rRZRβ (6,9,10) with the highest identity (89 and 91%) in the DBD. The DBD of RORγ exhibits a much lower homology with those of other nuclear receptors such as hRARα and mRev-erbβ (64 and 63% identity, respectively)(14,15,17). These results suggest that RORγ is a novel member of the ROR/RZR orphan receptor subfamily. High homology was observed between the T-box of RORγ and those of hRZRα/RORα, rRZRβ and Rev-erbβ (Fig. 2B). The T-box is a critical determinant in receptor binding to nucleotides extending 5' of the core-binding site PuGGTCA and may indicate that RORγ binds to similar response elements as Rev-erbβ and the other ROR's (Fig. 2B)(5,17).

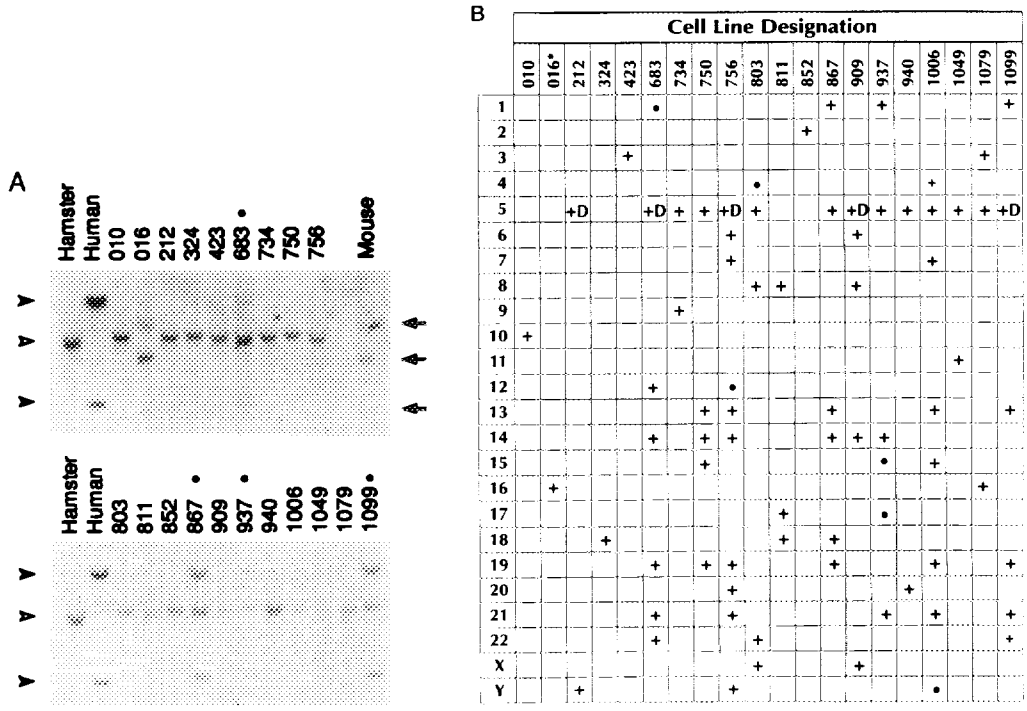

Fig. 3. Chromosomal localization of hRORγ. A. Chromosomal localization of human RORγ was determined by southern blot analysis using *Taq*I digested genomic DNA from different human/hamster or human/mouse hybrid cell lines. The blot was hybridized under stringent conditions with ^{32}P-labelled hRORγ probe (nt 1-250). Closed arrow heads indicate specific human DNA fragments; open arrow head, specific hamster DNA fragment; arrows, specific mouse DNA fragments hybridizing to the RORγ probe. B. Somatic cell hybrid panel specifying the human chromosome(s) present in the hybrid cell lines. +, indicates >30% of cells contain the given human chromosome; •, indicates 5-30% of cells contain given human chromosome; D, indicates the presence of multiple deletions in the respective human chromosome.

Chromosomal localization of RORγ - To determine on what human chromosome RORγ is located, *Taq*I-digested genomic DNA isolated from a panel of human/hamster and human/mouse hybrid cell lines was examined by Southern analysis with a ^{32}P-labeled hRORγ probe. The hRORγ probe hybridized only to human-specific DNA from the hybrid cell lines 683, 867, 937 and 1099 (indicated by ● in Fig. 3A). Fig. 3B shows that the hybrid cell line 683 contains human chromosome 1, 5, 12, 14, 19, 21 and 22, the cell line 867 contains human chromosome 1, 5, 13, 14, 18 and 19, 937 contains 1, 5, 14, 15, 17 and 21, and that the cell line 1099 contains 1, 5, 13, 19, 21, and

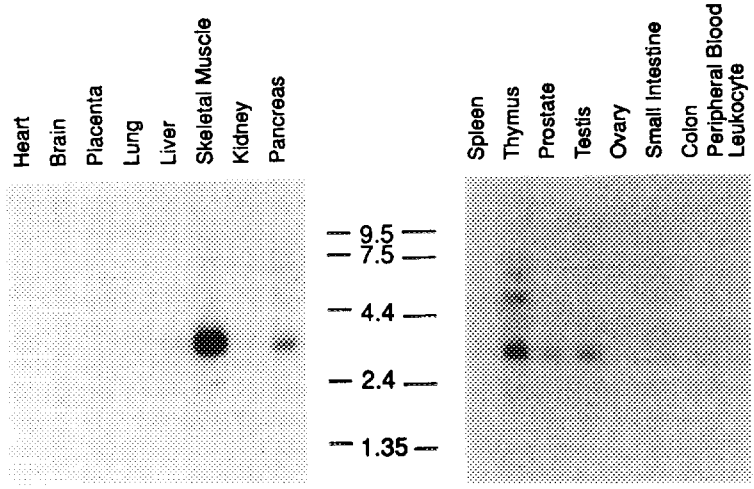

Fig. 4. Tissue distribution of RORγ mRNA. Poly(A)$^+$RNA (2μg) isolated from different human tissues was analyzed by Northern blot analysis using a [^{32}P]-labeled RORγ probe. RNA size markers are indicated between the two blots.

21. Since only chromosome 1 and 5 are common to these cell lines and in contrast to chromosome 5 no other cell line contained human chromosome 1 (Fig. 3B), it was concluded that the RORγ gene is located on human chromosome 1.

Tissue-specific expression of RORγ - To study the tissue-specific expression of RORγ mRNA, we performed Northern blot analysis on poly(A)$^+$ RNA prepared from a variety of human tissues (Fig.4). RORγ was expressed in several tissues, predominantly as a 3.2 kb mRNA. This RORγ mRNA was most highly expressed in skeletal muscle and at moderate levels in the thymus where the RORγ probe in addition to the 3.2 kb transcript hybridized to two other transcripts of 7.2 and 5.2 kb. RORγ was expressed at low levels in the pancreas, the tissue from which the original RORγ–DBD fragment was cloned, and in the prostate, testis, heart and liver. The size of the transcripts to which RORγ hybridizes as well as the tissue distribution of RORγ is different from those of RORα/RZRα and RZRβ (6,9,10). RORα/RZRα is expressed ubiquitously with highest abundance in peripheral blood leukocytes as transcripts of different size (mainly 15kb and 2.4kb)(9). RZRβ appears to be uniquely expressed in the brain as transcripts of 10kb and 2.4kb (10). These differences in tissue distribution suggest that each ROR/RZR receptor has a different function and regulates gene expression of different biological processes.

REFERENCES

1. Evans, R. M. (1988) Science, 240, 889-895.
2. Green, S., & Chambon, P. (1988) Trends. Genet., 4, 309-314.
3. Laudet, V., Hanni, C., Coll, J., Catzeflis, F., & Stehelin, D. (1992) EMBO J., 11, 1003-1013.
4. O'Malley, B. W., & Conneely, O. M. (1992) Mol. Endocrinol., 6, 1359-1361.
5. Wilson,T. E., Fahrner, T. J., & Milbrandt, J. (1993) Mol. Cell. Biol., 13, 5794-5804.
6. Giguere, V., Tini, M., Flock, G., Ong, E., Evans, R. M., & Otulakowski, G. (1994) Genes Dev., 8, 538-553.
7. Glass, C.K. (1994) Endocrine Reviews. 15, 391-407.
8. Giguere, V. (1994) Endocrine Reviews, 15, 61-795.
9. Becker-Andre, M., Andre, E., and DeLamarter, J. F. (1993) Biochem. Biophys. Res. Commun., 194, 1371-1379.
10. Carlberg, C., van Huijsduijnen, R. H., Staple, J. K., Delamarter, J. F., & Becker-Andre, M. (1994) Mol. Endocrinol., 8, 757-770.
11. Hirose, T., Fujimoto, W., Yamaai, T., Kim K. H., Matsuura, H. & Jetten, A. M. (1994) Mol. Endocrinol., 8, in press.
12. Hirose, T., O'Brien, D. A., & Jetten, A. M. (1994) Gene in press.
13. Hirose, T., Koga, M., Kouhara, H., Kishimoto, S., Matsumoto, K. & Sato, B. (1990) Cancer. Res., 50, 5060-5064.
14. Petkovich, M., Brand, N. J., Krust, A., and Chambon,P. (1987) Nature, 330, 444-450.
15. Giguere, V., Ong, E. S., Segui, P. and Evans, R. M. (1987) Nature, 330, 624-629.
16. Yu, V. C., Delsert, C. Anderson, B., Holloway J. M., Devary, O. V., Naar, A. M., Kim, S. Y., Boutin, J. M., Glass, C. K., & Rosenfeld, M. G.(1991) Cell, 67, 1251-1266.
17. Forman, B. C., Chen, J., Blumberg, B., Kliewer, S. A., Henshaw, R., Ong, E. S., and Evans, R. M. (1994) Mol. Endocrinol., 8, 1253-1261.
18. Mangelsdorf, D. J., Ong, E. S., Dyck, J. A., & Evans, R. M. (1990) Nature, 345, 224-229.