Endocrine, *vol. 8, no. 2, 123–134, April 1998* 0969–711X/98/8:123–134/$11.00

The Genomic Structure and Chromosomal Location of the Human TR2 Orphan Receptor, a Member of the Steroid Receptor Superfamily

Din-Lii Lin,[1,2] Shi-Qi Wu,[2] and Chawnshang Chang[1,2]

[1]George Whipple Laboratory for Cancer Research, Departments of Pathology, Urology, and Biochemistry, University of Rochester Medical Center, Rochester, NY; and [2]University of Wisconsin Comprehensive Cancer Center, University of Wisconsin, Madison, WI

Human TR2 orphan receptor, isolated from the testis and prostate, is a member of the steroid/thyroid hormone receptor superfamily. With the screening of a human genomic library and the combination of primer walking and PCR sequencing, we found that the entire TR2 orphan receptor gene coding region and 5'-untranslated region feature 13 introns and 14 exons, and that the consensus splice sequences (GT–AG) are present in all intron–exon boundaries. Within the region that codes for the DNA binding domain, TR2 orphan receptor gene has a distinct intron–exon junction. Whereas all other known steroid receptors have one splice site that separates their first and second zinc fingers in the DNA binding domain, TR2 orphan receptor has a rare splice site located in the middle of its first zinc finger. The identification of specific junction sequences for potential alternative splicing sites helps to explain the existence of multiple forms of TR2 orphan receptor cDNA (TR2-5, 7, 9, 11). The S1 nuclease protection assay for TR2 message revealed that there are multiple transcription initiations, and that the major cap site surrounded by an initiator-like sequence is located at the 104th nucleotide upstream from the translation start codon. Sequence analysis of a 2.7-kb DNA fragment upstream of the TR2 orphan receptor translation start codon unveiled several potential *cis*-acting elements, such as AP-1, HNF-5, GATA1 binding sites, and GC boxes. Using fluorescence *in situ* hybridization combined with a high-resolution G-banding technique, we found that the TR2 orphan receptor gene was mapped to human chromosome 12 at band q22, whereas the structurally closely related TR4 orphan receptor gene was mapped to human chromosome 3 at band q24.3.

Key Words: TR2; TR4; genomic structure; chromosomal location; transcription initiation.

Abbreviations: AR, androgen receptor; PR, progesterone receptor; RARβ, retinoic acid receptor β; CRBPII, cellular retinol binding protein II; DBD, DNA binding domain; LBD, ligand binding domain; RT-PCR, reverse transcriptase polymerase chain reaction; RXRβ; retinoid X receptor β; 5'-RACE, 5'-rapid amplification of cDNA end; GCTs, germ-cell tumors.

Introduction

Orphan receptors that bind to an unidentified ligand(s)/hormone(s) comprise the vast majority of the steroid/thyroid receptor superfamily. Members of this family are characterized by a highly conserved DNA binding domain (DBD) and a carboxy-terminal ligand binding domain (LBD). Most have been discovered by crosshybridization with authentic steroid receptor cDNAs *(1)*. The cDNAs of the human testicular orphan receptors, TR2 and TR4, share a high homology. They were isolated in the screening of human testis and prostate cDNA libraries using probes homologous to the highly conserved DBD of the steroid hormone receptor *(2,3)*. Four cDNA clones (TR2-5, 7, 9, and 11) were identified and found to share identical sequences in both the N-terminal domain and putative DBD, but differ in the length of LBD *(4)*.

The mechanism of action of the TR2 and TR4 orphan receptors remains largely unknown. However, there are two factors that may dictate how these two orphan receptors might function: (1) the existence of putative ligand/activator and (2) the cognate target genes that the orphan receptor influences. Our previous work *(5)* has shown that the TR2/AR/TR2 chimeric receptor, with the N-terminal and C-terminal domain of the TR2 orphan receptor, was not

Received October 10, 1997; Revised December 23, 1997; Accepted January 6, 1998.

Author to whom all correspondence and reprint requests should be addressed: Dr. Chawnshang Chang, George Whipple Laboratory for Cancer Research, Department of Pathology, Urology, and Biochemistry, University of Rochester Medical Center, 601 Elmwood Ave. Box 626, Rochester, NY 14642. E-mail: chang@pathology. rochester.edu

constitutively active when expressed in monkey kidney COS-1 and human prostate PC-3 cells in the absence of any added factors. In contrast, the human TR3 orphan receptor, another orphan receptor isolated in our laboratory *(6)*, was constitutively active in the same experimental system. Furthermore, a chimera receptor PR/PR/TR2 in which the N-terminal domain and DBD of the TR2 orphan receptor were replaced by those of the progesterone receptor, could be activated through a signal transduction pathway initiated at the cell membrane by the neurotransmitter dopamine *(7)*. These results suggested the necessity of a ligand/activator for the activation of the TR2 orphan receptor.

Among all adult tissues, testis is the only organ that expresses significant levels of TR2 *(8)*. However, Becker-Andre et al. have found that TR2 orphan receptors represent the major part of the steroid receptor mRNA pool amplified using total RNA isolated from human umbilical vein endothelial cells (HUVEC) *(9)*. *In situ* data on the TR2 orphan receptor gene expression pattern during mouse embryogenesis has also suggested that the TR2 orphan receptor gene may play an important role in diverse aspects of development, including organogenesis, proliferation, and differentiation of the neural system (Young, W. J., Lee, Y., Smith, S., and Chang, C., manuscript submitted). Collectively, these observations suggest that the TR2 orphan receptor gene is strongly and widely expressed throughout the entire animal development. In contrast, the TR4 orphan receptor has been detected with distinct abundance in brain regions, such as the hippocampus and cerebellum, suggesting that TR4 participates in neurogenesis *(3)*. Abundant TR4 transcripts are also detectable in the active proliferating zone of the developing nervous system of embryos *(10)*. This expression pattern is common for nuclear factors that function by either maintaining cells in an undifferentiated status or promoting cells into early differentiation.

Steroid/thyroid receptors bind to specific DNA sequences named hormone response elements (HREs) to regulate the expression of their target genes *(1)*. In our attempt to identify the TR2 and TR4 orphan receptor response elements, we found that the TR2 orphan receptor can bind to artificial hormone response elements (two half-sites of AGGTCA with a variety spacing in between) with different affinity: DR1 > DR2 > DR5, DR4, DR6 > DR3 *(11)*, whereas a similar tendency was observed in the case of TR4 (Lee, Y. and Chang, C., manuscript in preparation) as well. Based on this clue, we found that these orphan receptors may be capable of regulating the expression of a number of genes in ways such as repressing the RAR/RXR mediated induction of the RARβ or CRBPII gene *(11,12)*, enhancing the vitamin D-induced HRE 24-hydroxylase (Chang, C. et al., manuscript in preparation) and suppressing the expression of the simian virus 40 major late promoter *(13,14)* and the erythropoietin gene *(15)*.

In this article, we provide detailed information on the genomic structure of the TR2 orphan receptor and the chromosomal location for both TR2 and TR4 genes, which will assist in future studies of the TR2 and TR4 orphan receptors.

Results

Cloning and Sequencing of the Intron–Exon Junctions of the Human TR2 Orphan Receptor Gene

We screened a human placenta genomic λ FIX II library (Stratagene) with ^{32}P-labeled human TR2 orphan receptor cDNA fragments representing the N- and C-terminal regions of the receptor. Seven positive clones were isolated with overlapping sequences covering the entire coding region and 5'-promoter region. A comparison of the partial genomic sequences obtained from the combination of primer walking and PCR sequencing with the cDNA sequence allowed the identification of 13 introns and 14 exons in these regions (Fig. 1 and Table 1). All the intron–exon junction sequences follow the GT–AG splicing rule (Table 1). One clone that carried an insertion of approx 20 kb from the 5'-end of the gene was subjected to further characterization. According to the restriction digestion pattern and Southern blot analysis, this genomic insert contains at least 2.7 kb of the DNA fragment upstream from the first intron of the TR2 gene. Additionally, we also cloned a library candidate that codes for the DBD region of the TR4 orphan receptor.

The first exon of the TR2 gene, 97 bp in length, contains 96% of the entire 5'-UTR which is 104 bp in length. The second and third exons, 61 and 231 bp, cover most of the N-terminal domain. The fourth exon, 79 bp in length, encodes the rest of the N-terminal domain and half of the first zinc finger; the fifth exon, having 180 bp, encodes the rest of the DBD. For the C-terminal end junction of this exon of both the TR2 and TR4 (data not shown) genes, the position is well conserved without exception among all of the steroid hormone receptors published thus far *(16)*. For the N-terminal end, the splicing sites of the TR2 and TR4 (data not shown) genes are, again, exactly matched. However, whereas all other known steroid receptors have one splice site that separates the first and second zinc fingers of their DBD, the TR2 and TR4 genes each have this distinctive splice site located in the middle of its first zinc finger (Fig. 2). This is strikingly similar to the genomic structure of the retinoid X receptor (RXR) β gene *(17)*. The sixth exon, covering 148 bp, contains the hinge domain between the DBD and the LBD. The rest of the exons (7th–14th) encode the entire LBD of the TR2 orphan receptor.

According to sequence comparison among the TR2-5, 9, and 11 receptor isoform cDNAs and genomic structure data (Fig. 1), we found that their sequences were identical from the 5'-end until the point of junction K. Using two specific primers, TR2-5-u and TR2-9-u, we discovered that the positions of their specific intron acceptor sites were well matched (Fig. 3B). Moreover, by employing common "d" primer plus isoform-specific primers (TR2-11-u, TR2-5-u,

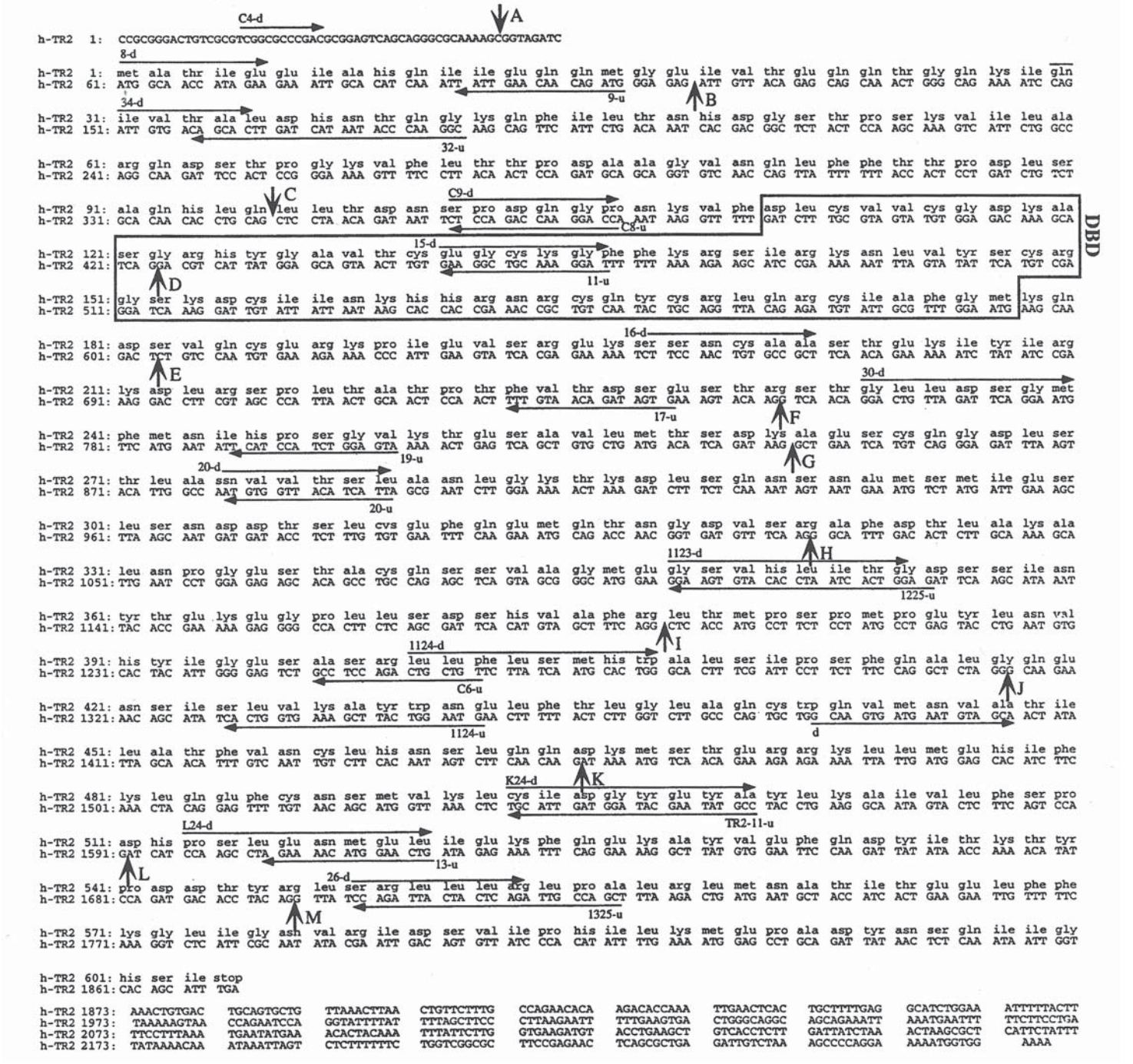

Fig. 1. The distribution of intron insertion sites along the TR2 cDNA sequence. Arrows and capital letters denote intron insertion sites. The DBD is boxed. The position of the oligomers used for PCR sequencing is illustrated.

and TR2-9-u, respectively) to PCR-amplify each individual intron fragment, we found that the resulting sizes of the introns were 2.6 kbp (TR2-11 specific, Fig. 3A, lane 2), 620 bp (TR2-5 specific, Fig. 3A, lane 3), and 860 bp (TR2-9 specific, Fig. 3A, lane 4). These data suggest that the isoform-specific cDNA sequences for TR2-5 and TR2-9 were part of intron K, and the TR2-5, TR2-9, and TR2-11 cDNAs were derived owing to alternative splicing sites within intron K (Fig. 3B).

Characterizing the 5'-Flanking Region

The upstream 2.7-kb 5'-flanking region of the gene was subcloned and sequenced. The complete nucleotide sequence (GenBank accession number U19026) is shown in Fig. 4. A sequence homology search of the database for the transcriptional factor binding motif revealed several potential *cis*-acting elements in this promoter region (Fig. 4). Two GC boxes, known as the putative SP1 binding site *(18)*, were located near the cap site. Four AP-1 motifs

Table 1
Intron–Exon Junction Sequences of TR2-11 Orphan Receptor Gene

Intron	Exon	Intron sequences		Exon
A	CAAAAGCG	GTAAGAAACGCTGGGTGGGGACACCAGCGACCGCGT	GAGTTTAATTATAAAACATCTTTTTATTTTTGGCAG	G TAG ATC
B	ATG GGA GAG met gly glu 16 17 18	GTATGTCTAGTTTACGGTTGATGCACTGGCCCACCA	TTGAGTTTCATTGATTATTTTTCATGATTTTGTTAG	ATT GTT ACA ile val thr 19 20 21
C	CAC CTG CAG his leu gln 93 94 95	GTAAATAACTATAGAAGTGTCTTTTCCTCTTACATA	TTTGTTTTCTTTTTCTCTTTTTTTTTTTTTTTTTAG	CTC CTA ACA leu leu thr 96 97 98
D	CGA TCA G ala ser g 120 121 1	GTAACATTAAGAATAACCTCCTAAAAACAGAAAAAA	GTGACTTTTATATTTTCTTAAACATTTTCTCTCTAG	GA CGT CAT ly arg his 22 123 124
E	CAA GAC T gln asp s 180 181 1	GTATGTATTAGCTTTTAAGGAGAAAATACTTTTTA	TATATTGTGATATAATAGTTTAACACATTTTTGTAG	CT GTC CAA er val gln 82 183 184
F	AGT ACA AG ser thr ar 229 230 23	GTAAAGTAAAACAGACACTGAGTAGCTGGTCATTTT	GAAACTCAGGCCAAAATTTGTGCTGATTTGTTACAG	G TCA ACA g ser thr 1 232 233
G	TCA GAT AAG ser asp lys 259 260 261	GTGTGTTTAGTACAATTTAAAGCTTAGATAATTTTA	CCAATTCACAGAATTTTAATACCCTTTTTTTTAAAG	GCT GAA TCA ala glu ser 262 263 264
H	GTT TCA AG val ser ar 320 321 32	GTAAGGTATCTTTTGATTATCAATGTCCTTTTGTTT	TGTTGTCAAGCTCATAGCTTCTGATGTCTCATACAG	G GCA TTT g ala phe 2 323 324
I	GCT TTC AGG ala phe arg 377 378 379	GTATTTCTATATGTTTTTTTTAACTTGTATGTTGTG	TCATAAGATAATTTACCTAAATCCCTGTCATATTAG	CTC ACC ATG leu thr met 380 381 382
J	GCT CTA GG ala leu gl 416 417 41	GTGAGTGTTTTGAAATCTTGAATATAAATGTGATTT	TAATCTGATTTGTGCTTTTATTACTGTTTTTTAATAG	G CAA GAA y gln glu 8 419 420
K	CAA CAA G gln gln a 463 464 4	GTAAAAAGCACCTTCCTTACCTTTTCCAGGTTGTAG	CAAACACTAGAAGCATTTGTTTATTTTTAATTAAAG	AT AAA ATG sp lys met 65 466 467
L	AGT CCA G ser pro a 509 510 5	GTATATAAATTACTTTTATCAAATAAATGTTCACGA	GAACCTAGGGTACCAATGTTTTCTTTAATGTTTCAG	AT CAT CCA sp his pro 11 512 513
M	ACC TAC AG thr tyr ar 544 545 54	GTATCTAGAGTTAAATTCTATATAATTAAAATAATT	CCCTTGTTCCTCTCCCTTTTTTCTTTCTTTTTTTAG	G TTA TCC g leu ser 6 547 548

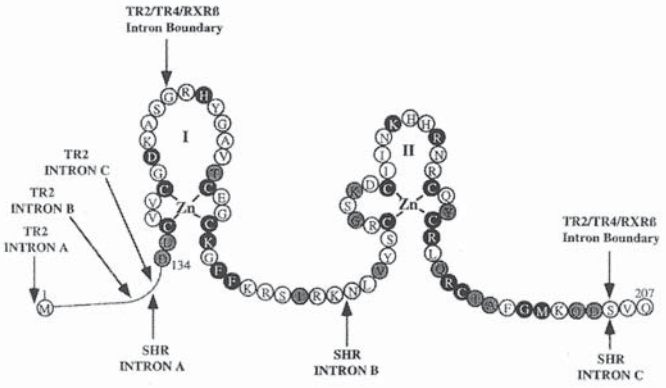

Fig. 2.

(19), two AP-2 binding site-like sequences *(18)*, and one AP-3 binding site-like sequence *(18)* were found along the 5'-flanking region. Among them, two AP-1 (at –374/–368 and –355/–349) and one AP-2 (at –124/–115) consensus motifs resided within a close range. Noticeably, this spe-

Fig. 2. *(left)* The Zn-fingers of human TR2 orphan receptor. Comparison of the intron positions in and around the zinc fingers of the TR2 orphan receptor and steroid hormone receptors. The identical amino acid residues in all known steroid receptors are depicted in black circles, whereas the distinctive amino acid residues between TR2 and TR4 orphan receptors are depicted in gray circles. The positions of the introns in the TR2/RXRβ orphan receptors relative to common splicing sites for other known steroid hormone receptors (SHR) are indicated.

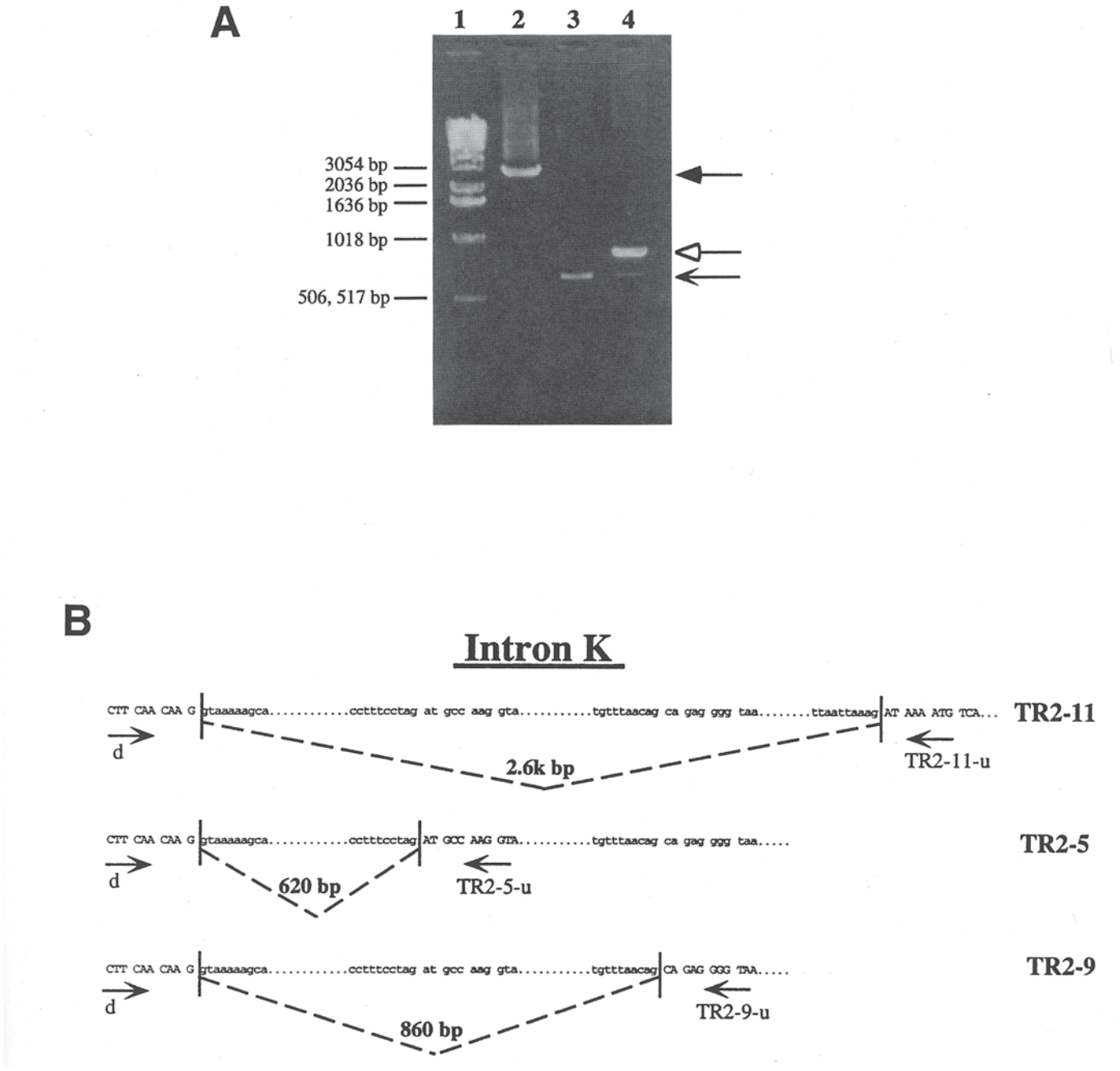

Fig. 3. The TR2-unique and isoform-specific junction of the TR2 orphan receptor gene. **(A)** Result of isoform-specific PCR fragments. Lane 1 is a 1-kb marker (Gibco BRL); lane 2 is a TR2-11-specific fragment (2.6k bp); lane 3 is a TR2-5-specific fragment (620 bp); lane 4 is a TR2-9-specific fragment (860 bp). **(B)** Alternative splicing sites within intron K for TR2 isoforms: TR2-5, TR2-9, and TR2-11. DNA sequences shown in upper case represent exonic sequences and in lower case represent intronic sequences. Arrows indicate the primers employed for PCR amplification of the intronic fragments.

cific AP-2 motif appears to overlap with one of the GC boxes (–119/–114). Other potential *cis*-acting elements found harbored in the promoter region included the CArG core sequence required for muscle-specific transcriptional stimulation *(19)*, the C/EBP motif bound by the liver-abundant activator *(18)*, CP2-γ-FBG *(20)*, the cAMP-responsive EivF/CRE *(21)*, the erythroid cell-specific GATA-1 *(18)*, and several other response elements described therein *(18)*.

Determination of Cap Site in the 5'-Flanking Region

An upward primer, named TR2-S1-u: 5'-GCTGCCACTCGT GGATTGGG-3', starting at the 61st nucleotide from the translation start codon, was used to generate a single-strand DNA probe, 172 nucleotides in length, to anneal to RNA from three different cell lines: HeLa, HepG2, and LNCaP. A DNA-sequencing reaction using the same primer on a genomic clone was included in the same gel to align the cap

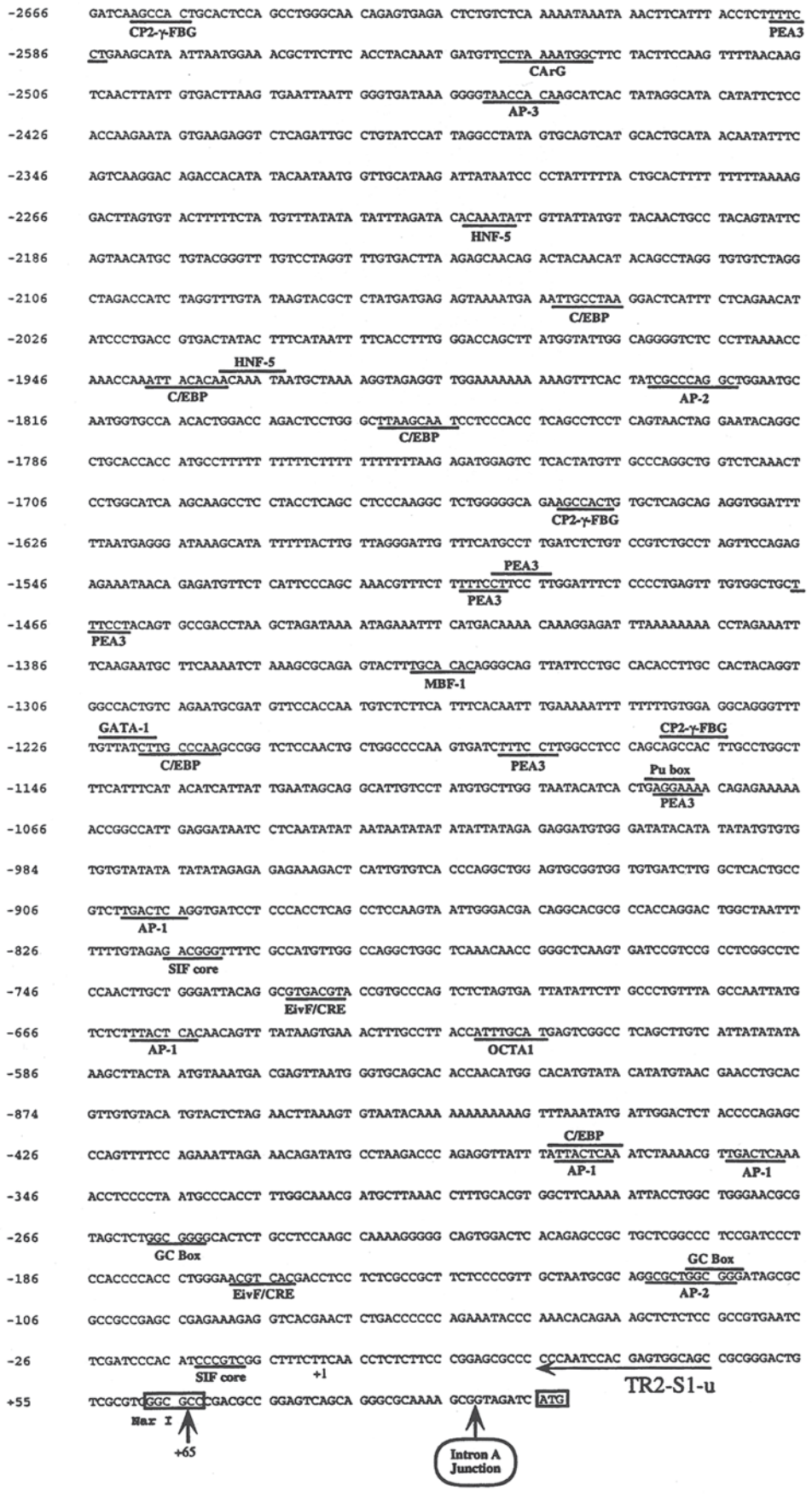

Fig. 4. Sequence of the human TR2 orphan receptor 5'-flanking region from –2666 to +107 (ATG). "+1" indicates the major TIS. Potential *trans*-factor binding sites are marked. The junction position for intron A is indicated. Restriction *Nar*I site used to ligate 5'-flanking sequence is shown. (GenBank assigned accession number for this sequence is U19026.)

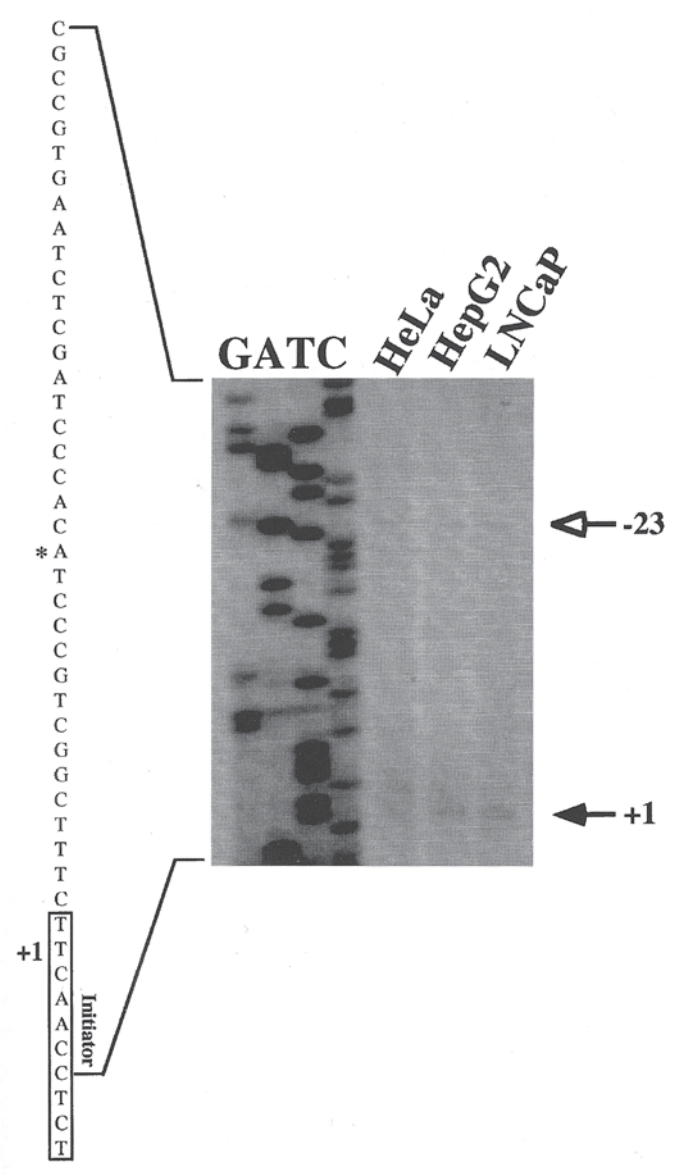

Fig. 5. Mapping of the cap site by the S1 nuclease protection analysis. A single-stranded DNA, 172 nucleotides in length, extended from the primer TR2-S1-u to the *Fsp*I restriction site, was hybridized to total RNA isolated from HeLa, HepG2, and LNcaP cells, and extended as described in Materials and Methods. The "+1" indicates the major cap site and asterisks indicate the multiple TIS. The initiator-like sequence is boxed.

site of mRNA to the genomic DNA sequence. The result displayed multiple transcription initiations (Fig. 5); the major and the minor cap sites were located at the 104th and 127th nucleotide (designated as +1 and –23) upstream from the translation start codon, respectively. These data matched well with the results we obtained from 5'-RACE assay using rTth polymerase (Clontech, Palo Alto, CA) (data not shown). There is no TATA box found neighboring the transcription initiation by DNA sequence comparison. In contrast, an initiator-like sequence (5'-TTCAACCTCT-3') was identified in the sequences surrounding the major cap site *(22)* (Fig. 5).

Chromosome Mapping

To find out whether the two highly similar receptors are also close in chromosomal location, we set forth to map their positions. Thirty-six metaphase spreads were used to analyze the presence of FITC-labeled fluorescent signals. Specific hybridization spots on the long arm of chromosome 12 were found in 32 metaphases. The majority of these metaphases (>75%) revealed two signals on each chromosome (one signal on each chromatid) (Fig. 6A). High-resolution G-banding of chromosome 12s and the same chromosomes hybridized with FITC-labeled TR2 orphan receptor gene were compared to determine the location of the hybridization signals. The gene was mapped to chromosome 12, at band q22 (Fig. 6B). A schematic diagram of the high-resolution G-banded chromosome 12 (550 and 850 bands) shows the precise location of the TR2 orphan receptor gene on chromosome 12 as well as several other contiguous genes in this region (Fig. 6C) *(23)*. For TR4 gene mapping, a total of 42 metaphase spreads were used to analyze the presence of hybridization signals. Specific hybridization spots on the short arm of chromosome 3 were found in 36 metaphases. The majority of these metaphases (>70%) revealed two signals on each chromosome (one signal on each chromatid) (Fig. 7A). G-banded chromosome 3s and the same chromosomes hybridized with FITC-labeled TR4 gene as well as rhodamine-labeled 3p telomeric probe were compared to determine the location of the hybridization signals. As shown in Fig. 7B, the hybridized signals (arrowheads) were mapped to chromosome 3, at band p24.3. A schematic diagram of G-banded chromosome 3p (550 band level) indicates the precise location of the TR4 orphan receptor gene on the short arm of chromosome 3 (Fig. 7C).

Discussion

The information on intron–exon junction sites clearly is very helpful for explaining the possibility of the existence of multiple forms of the TR2 orphan receptor (TR2-5, TR2-7, TR2-9, TR2-11) identified from the testis and prostate cDNA libraries. On aligning the cDNA sequence, one will see that TR2-7 has the same sequence as TR2-5, with an addition of 429 bp at the 3' connecting to the DBD, which derived an early terminated protein isoform with the loss of the putative LBD. TR2-5, TR2-9, and TR2-11 carry the identical sequence of the N-terminal and DBD with a divergent C-terminal sequence. Observations from Northern blot results *(24)* suggest that TR2-11 most likely represents the major form of the TR2 orphan receptor gene.

Unexpectedly, based on the genomic analysis result reported here, we found that the DNA sequence of the entire intron E, a 429-bp DNA fragment, matches exactly the extra sequence previously identified at TR2-7 cDNA *(4)*. Moreover, a short transcript of the SpSHR2 orphan receptor, a sea urchin relative of the TR2 orphan receptor, has also

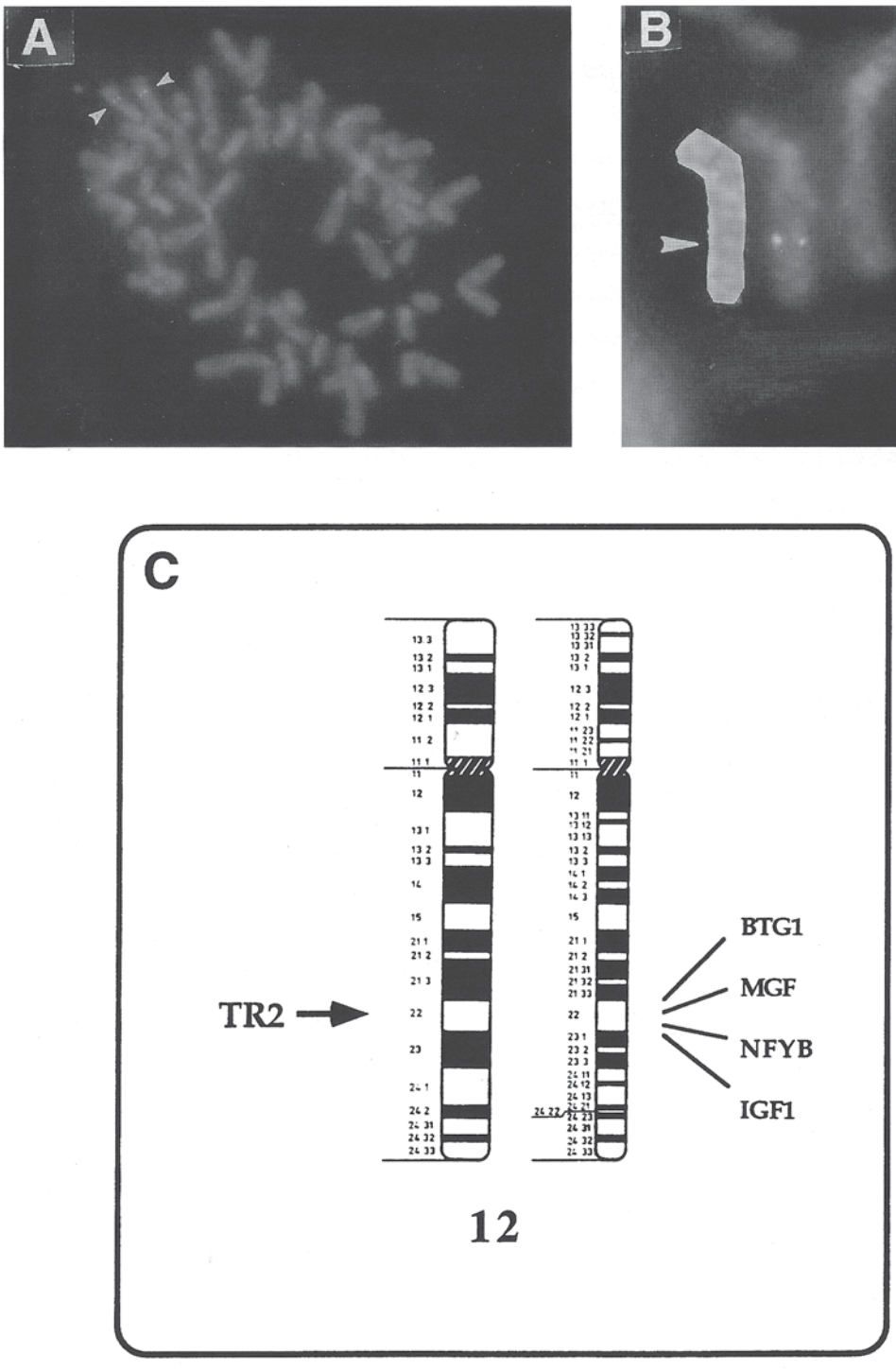

Fig. 6. The chromosomal location of TR2 gene is 12q22. **(A)** Chromosome mapping. FISH analysis of human metaphase showing the TR2 orphan receptor gene signals on the long arms of both chromosomes 12s (arrowheads). **(B)** High-resolution G-band of chromosome 12. Partial karyotype of high-resolution G-banded chromosome 12 is shown along with the same chromosome hybridized with the FITC-labeled TR2 orphan receptor gene, indicating that the signals are mapped at 12q22. **(C)** Ideogram of the high-resolution G-banded chromosome 12. Ideogram of chromosome 12 (550- and 850-band levels) shows the TR2 orphan receptor gene location and four contiguous genes mapped at the same region.

been shown to skip its putative LBD by Kontrogianni-Konstantopoulos et al. *(25)*. These data strongly suggest that TR2-7 is a cDNA derived from alternative splicing. Assuming that there is a specific ligand for the TR2 orphan receptor, TR2-7 should not be able to bind to it. It is possible, therefore, that regulatory functions attributed to these truncated receptors may be so important as to be conserved from sea urchin to mammals.

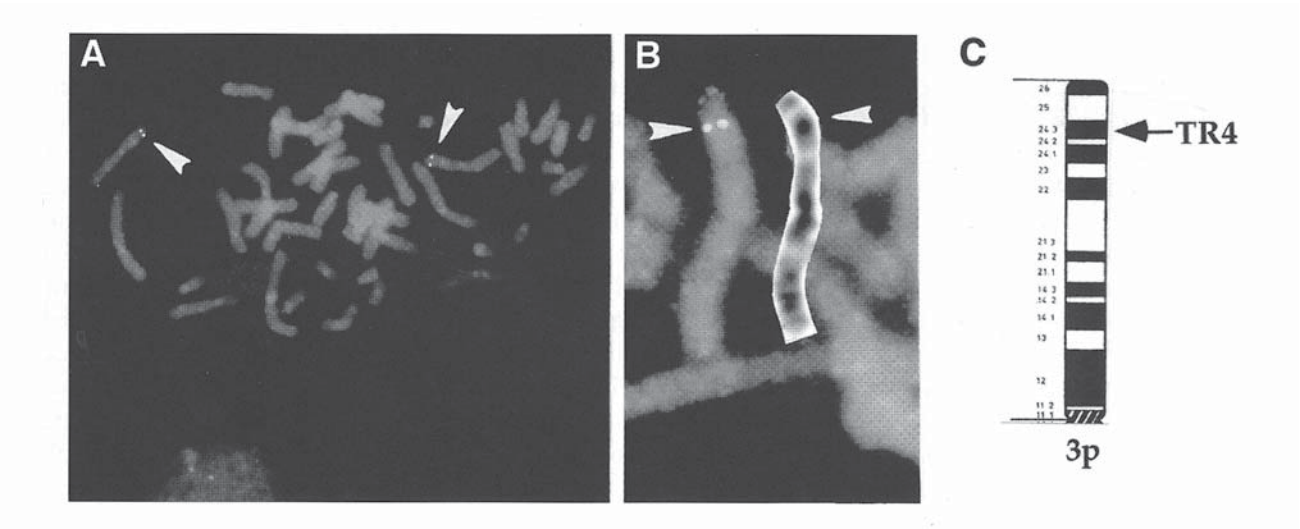

Fig. 7. The chromosomal location of TR4 gene is 3p24.3. **(A)** Chromosome mapping. FISH analysis of human metaphase spreads showing the TR4 orphan receptor gene signals on the short arms of both chromosomes 3s (arrow heads). **(B)** Dual-color FISH and G-banded chromosome 3. Partial karyotype of G-banded chromosome 3 is shown along with the same chromosome cohybridized with the FITC-labeled TR4 orphan receptor gene and the rhodamine-labeled 3p telomeric probe, indicating the signals are mapped at 3p24.3 (arrowheads). **(C)** Ideogram of the high-resolution G-banded chromosome 3p. Ideogram of the high-resolution G-banded chromosome 3p (550-band level) shows the TR4 orphan receptor gene location (arrow).

Furthermore, we found that the sequence differences among TR2-5, TR2-9, and TR2-11 appear to diverge right after the intron–exon junction K. This suggests that TR2-5 and TR2-9 may be derived from splicing alternative sequences. Indeed, the alternative splicing form of a receptor exists in other members of the steroid receptor superfamily, such as thyroid hormone and glucocorticoid receptors *(26,27)*. These alternative forms of receptors either share a redundant function or behave in a mutually exclusive manner. Therefore, the multiple forms of TR2 orphan receptors may increase the complexity and potential importance of the biological functions of this gene. The *cis*-controlling element for the regulation of the alternative splicing should lie within intron K as well. Future investigation into this region would be necessary to reveal the control mechanism for differential splicing.

Structural analysis of the TR2 orphan receptor gene revealed that its intron–exon structure is similar to that of other members of the steroid receptor superfamily with some notable exceptions: whereas all other known steroid receptors have their N-terminal domain encoded by one exon *(16)*, the TR2 orphan receptor gene has four exons to make up the entire N-terminal domain (Fig. 2). Furthermore, we found that the genomic structure of the TR2 and TR4 (data not shown) orphan receptor genes appeared to be highly similar to that of the RXRβ gene in the region that codes for the DBD. Unlike other known steroid receptor gene families that harbor an intron to separate two zinc fingers, the related intron of TR2/RXRβ genes is located at the center of the first zinc finger (Fig. 2). Conversely, the next intron is located at the C-terminal of the DBD and is exclusively conserved within all receptor genes. These data strongly suggest that all members of the steroid receptor superfamily may have evolved from one ancestral gene. Moreover, RXR is a mammalian counterpart of the *Drosophila* ultraspiracle (USP) which functions as the partner of the ecdysone receptor (EcR) *(28)* and the rare intron junction located in the middle of the first zinc finger has also been shown to be conserved in the TR2 sea urchin relative, SpSHR2 gene *(25)*.

The S1 nuclease protection assay result (Fig. 5) revealed multiple transcription-initiation bands and the major cap site located at the 104th nucleotide upstream of the translation start codon. To date, relatively little is known about the transcriptional regulatory processes implicated in controlling levels of TR2 orphan receptor mRNA. Current evidence indicates that the TR2 orphan receptor can be expressed in many tissues during embryogenesis (Chang, C. et al., manuscript in preparation). Further analysis of the 5'-flanking regions of the human TR2 orphan receptor gene may provide potential insight into the basis for the control of this gene. Based on our data (Fig. 4), the surrounding sequences near initiator-like sequences are GC-rich and do not contain a consensus TATAA motif. Transcription by RNA polymerase II commonly utilizes the Cys^2-His^2 zinc finger protein SP1 in this type of promoter *(29)*. This 105-kDa sequence-specific transcription factor binds to guanine-rich binding sites of TATA-less promoters in a wide variety of constitutively expressed cellular genes, including the human low-density lipoprotein receptor *(30)*, retinoblastoma gene product *(31)*, and insulin receptor *(32)*. Two SP1-like (GC box binding protein) sites were also found in the TR2 orphan receptor gene 5'-flanking region at –256 and –116. It has been demonstrated that SP1, initiator, TFIID, and undetermined coactivators are required for transcription from some TATA-less promoters *(33)*. Whether

SP1 sites within this regulatory region play any significant role in the expression of the TR2 orphan receptor gene remains to be determined.

We reported earlier that the expression of the TR2 orphan receptor gene may be repressed by androgen *(2,24)*. Northern blot analyses demonstrated that castration could increase the TR2 orphan receptor mRNA levels and that the addition of androgen could reverse this increase in the rat ventral prostate and human prostate LNCaP cells. However, we were unable to locate any androgen response element-like sequence along the 2.7-kb 5'-flanking region. Perhaps additional upstream elements beyond these 2.7-kb 5'-flanking sequences or intron sequences are required for androgen to repress the TR2 orphan receptor expression.

More recently we demonstrated evidence suggesting that the tumor suppressor p53 may be the mediator for the radiation repression of TR2 orphan receptor expression *(34)*. The observation that there exists a similar expression pattern between p53 and the TR2 orphan receptor in murine spermatogenesis (Young, W. J. et al., paper submitted) may support the suggestion that the TR2 orphan receptor plays important roles in the p53 signal pathway. As a nuclear phosphoprotein, p53 can regulate target genes by direct or indirect interaction with DNA sequences *(35)*. Recent mapping of the p53-response elements (p53RE) in several p53 target genes also suggests that the p53RE may contain two copies of the 10-nt motif (RRRCWWGYYY) *(36)*. Since one copy of the motif is not sufficient to bind to the p53 tetramers, one might find that the only potential p53RE identified at nt –1697 in the human TR2 orphan receptor gene 5'-flanking region may not be enough for p53 to modulate TR2 orphan receptor gene expression. The repression of the TR2 orphan receptor by p53 may, therefore, function via other unknown mechanisms *(34)*.

We have determined that the TR2 orphan receptor gene is located on human chromosome 12q22. Interestingly, another orphan receptor, TR3, is also located in the long arm of chromosome 12 *(6)*. The TR3 orphan receptor gene has been mapped to the more proximal part of 12q at band 13.1, and the TR2 orphan receptor gene has been mapped to the distal part. Within this region reside several important genes, including B-cell translocation gene 1 (BTG1), mast cell growth factor (MGF), transcription factor NF-Y, B-subunit (NFYB), and insulin-like growth factor-1 (IGF-1) (Fig. 6C). Analysis of loss of constitutional heterozygosity of human male germ-cell tumors (GCTs) resulting from malignant transformation of premeiotic or early meiotic germ cells revealed two regions with frequent loss (>40%), one at 12q13 and the other at 12q22 (where the TR3 and TR2 genes are located, respectively). Deletions at those two regions characterized in GCTs may lead to the loss of some potential tumor-suppressor genes whose products regulate the normal proliferation of the spermatogonial stem cells *(37,38)*. Chromosomal abnormalities in the region of these genes have also been associated with specific forms of neoplasia *(39)*. Whether or not there is a potential linkage between the TR2 orphan receptor with other genes located in the neighboring area will be an interesting question to investigate.

The hybridized signal of the TR4 gene, on the other hand, was mapped to chromosome 3p24.3. It is intriguing to learn that the human placental thyroid hormone receptor β-gene (THRβ) or ERBA2 gene is also located at 3p24.3 *(40)*. Mutations have been frequently found in the hormone binding domain of this gene and resulted in an inherited generalized resistance to thyroid hormone *(41)*. Another gene that is located close to the TR4 locus is the RARβ gene, one of the members of the nuclear receptor superfamily. The product of this gene, which shares high homology with THRβ, is predominantly present in epithelial tissues and a novel ligand-responsive regulatory protein whose inappropriate expression in liver is associated with hepatocellular carcinogenesis *(42)*. The distinct chromosomal location shown here for the TR4 gene rules out the possibility that it may also be derived from part of the TR2 gene, although it is highly homologous in cDNA sequence to the TR2 gene. This finding suggests that the TR2 and TR4 orphan receptors, which are closely related but have two different identities, may constitute a subfamily in the steroid receptor superfamily. In sum, our data in this article provide valuable information that is helpful toward the understanding of TR2 orphan receptor.

Materials and Methods

Cloning and Sequence Analyses

A human placenta genomic λ FIX II library (Stratagene) was screened using a ^{32}P-labeled 320-bp *Eco*RI-*Pst*I fragment and a 1.3-kbp *Bgl*II-*Xho*I fragment of the human TR2 orphan receptor cDNA representing the N- and C-terminal regions of the receptor, respectively. Similarly, the same library (Stratagene) was screened using a ^{32}P-labeled 625-bp *Eco*RI-*Sca*I fragment and a 700-bp *Nae*I-*Eco*RI fragment of the human TR4 orphan receptor cDNA N- and partial C-terminal as probe. The membranes with 4×10^5 phage plaques were hybridized for 48 h at 42°C in hybridization buffer (50% formamide, 5X SSC, 5X Denhart's solution, 0.5% sodium dodecyl sulfate [SDS], 50 μg/mL heat-denatured salmon sperm DNA), and washed for 60 min with 2X SSC 0.1% SDS at room temperature, and for 30 min with 0.2X SSC 0.1% SDS at 55°C twice. Positive phage clones were identified through two consecutive screenings, and enriched and isolated by CsCl banding. DNA of positive phages was purified following the protocol described in *Molecular Cloning (43)*. These DNA clones then served as templates for PCR sequencing to obtain partial sequences by employing Thermo-Sequenase (Amersham, Cleveland, OH). Briefly, the cycle for PCR was as follows: denaturation at 94°C for 1 min; then, cycling, denaturation at 94°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C

for 1 min for 30 cycles; with one final additional extension at 72°C for 1 min. DNA sequencing was resolved in a glycerol-tolerant denaturing sequencing gel (Amersham). The sequences of the primers used for sequencing were as follows:

C4-d: 5'-CGCGTCGGCGCCCGA-3'
8-d : 5'-ATGGCAACCATAGAAGAA-3'
9-u: 5'-CATCTGTTGTTCAATA-3'
34-d: 5'-CAGATTGTGACAGCAC-3'
32-u: 5'-GCCTTGGGTATTATGATCAAGTGCTG-3'
C9-d: 5'-CTCCAGACCAAGGACC-3'
C8-u: 5'-GGTCCTTGGTCTGGAG-3'
15-d: 5'-GAAGGCTGCAAAGGAT-3'
11-u: 5'-ATCCTTTGCAGCCTTC-3'
16-d: 5'-TTCCAACTGTGCCGCT-3'
17-u: 5'-CACTATCTGTTACAAA-3'
30-d: 5'-GGACTGTTAGATTCAGGAATG-3'
19-u: 5'-TACTCCAGATGGATGA-3'
20-d: 5'-ATGTGGTTACATCATT-3'
20-u: 5'-AATGATGTAACCACAT-3'
1123-d: 5'-GGAAGTGTACACCTAATCACTGG-3'
1225-u: 5'-CTCCAGTGATTAGGTGTACACTTCC-3'
1124-d: 5'-CTGCTGTTCTTATCAATGCACTGG-3'
C6-u: 5'-ACAGCAGTCTGGAGGC-3'
d: 5'-GCAAGTGATGAATGTAGCA-3'
1124-u: 5'-CATTCCAGTAAGCTTTCACCAGTG-3'
K24-d: 5'-TGCATTGATGGATACGAATATGCC-3'
TR2-11-u: 5'-GGCATATTCGTATCCATCAATGCA-3'
L24-d: 5'-CCAAGCCTAGAAAACATGGAACTG-3'
13-u: 5'-CAGAACCATGTTTTCT-3'
26-d: 5'-CCAGATTACTACTCAGA-3'
1325-u: 5'-GCTGGCAATCTGAGTAGTAATCTGG-3'
TR2-5-u: 5'-CACTGTTTTGGGTAGTCATAAGCC-3'
TR2-9-u: 5'-GTAAGATGGGAATGTGATGCAGCC-3'

The clone that carries the 20-kb 5'-most of the genomic insert was digested as described below and subcloned into the respective restriction sites of pBluescript II SK(-) (Stratagene): *Not*I-*Eco*RI 6.2-kb, *Eco*RI-*Eco*RI 1.4-kb, *Eco*RI-*Eco*RI 0.7-kb, *Eco*RI-*Eco*RI 6-kb, *Eco*RI-*Eco*RI 0.9-kb, *Eco*RI-*Sal*I 5-kb. The 6.2-kb fragment was further cut with *Nar*I to remove the 5'-portion of the first intron. DNA sequences were analyzed using UWGCG (University of Wisconsin Genetics Computer Group) sequencing software. The complete upstream 2.7-kb 5'-flanking nucleotide sequence was deposited into the GenBank database under accession number U19026.

S1 Nuclease Protection Assay

The 5'-UTR primer, TR2-S1-u: 5'-GCTGCCACTCGT GGATTGGG-3', starting at the 61st nucleotide from the translation start codon was ^{32}P-end-labeled and served as a primer for PCR (same cycling conditions as mentioned above) generation of a single-strand DNA probe, 172 nucleotides in length, from a 2.7-kb fragment digested by *Fsp*I. This probe was then hybridized to total RNA from the TR2 positive cell lines in buffer from the S1 Assay kit according to the commercial protocol (Ambion, Austin, TX). The mixture was incubated at the suggested temperature for 1 h, precipitated, resuspended, and denatured, prior to analysis, on an 8% sequencing gel. A DNA-sequencing reaction using the same primer on a genomic clone was included in the same gel to align the initiation site of mRNA with the genomic DNA sequence.

Metaphase Cell Preparation

Metaphase spreads were obtained from PHA-stimulated lymphocytes of human peripheral blood. The harvesting of lymphocytes and the preparation of slides were carried out as described *(44)*. G-banding was performed after the air-dried slides were 1 wk old. The banded metaphase chromosomes were examined and photographed with an MAX-BX 40 Olympus microscope using a UWplan FL 100× objective (dry lens). The coordinates of each metaphase were recorded. The slides were then destained with two changes of 3:1 methanol/acetic acid for 5 min each, and air-dried for further use in fluorescence *in situ* hybridization (FISH).

Labeling of the DNA Probe

The probes used for localization of the human TR2 orphan receptor gene were plasmids p171C and p171D containing 7- and 6.4-kb *Sac*I genomic fragments covering genomic regions upstream and downstream from the DBD, respectively. Similarly, the probe used for localization of the human TR4 orphan receptor gene was plasmid PB7KhTR4 containing 7-kb *Xba*I-*Xba*I genomic fragments covering N-terminal coding regions of the gene. Probes were labeled by nick translation using biotin-14-dATP according to the manufacturer's instructions (Gibco BRL, Grand Island, NY). Unincorporated nucleotides were removed by repeated ethanol precipitation. The size of the probe fragments was between 300 and 600 nucleotides as assessed by 1.2% agarose-gel electrophoresis. The labeled fragments were stored at –20°C.

Fluorescence In Situ *Hybridization (FISH)*

The procedure used for FISH is a modification of the method described previously *(45)*. Chromosomal DNA was denatured by immersion of the slides in 70% formamide/2X SSC, pH 7.0 at 70°C for 2 min, followed by dehydration in an ethanol series, and air-dried. Prior to hybridization, human Cot-1 DNA (final concentration, 150 μg/mL) and sonicated herring sperm DNA (final concentration, 50 μg/mL) were added to the hybridization mixture (50% formamide/2X SSC/10% dextran sulfate, pH 7.0) containing the labeled probes (5–15 ng/μL). The probe hybridization mixture was denatured at 80°C for 10 min and cooled on ice immediately. Ten microliters were applied to each slide, the slides were covered with glass cover slips (22×22 mm^2), sealed with rubber cement, and then transferred to a humidified chamber. The slides were incubated at 37°C for 24 h. After

hybridization, the slides were washed in 1X SSC at 70°C for 5 min and soaked in 1X PBD at room temperature for 2 min. The probe was detected with 15 µL FITC-avidin at 37°C for 15 min, followed by 1X PBD washes (3 times, 2 min each). The fluorochrome signal was amplified by incubation with 15 µL antiavidin at 37°C for 15 min, followed by 1X PBD washes (three times, 2 min each), a second incubation with 15 µL FITC-avidin at 37°C for 15 min, and a final PBD wash. Metaphase cells were counterstained with 15 µL PI (0.6 µg/mL) at room temperature for 5 min. The slides were examined with the Olympus BX 40 fluorescence microscope equipped with an Olympus WIB filter.

Dual-color FISH was performed by cohybridization of biotin-labeled TR4 probe and digoxygenin-labeled human chromosome 3p telomeric probe (Oncor, Gaithersburg, MD). The hybridized probes were detected using FITC-conjugated avidin and rhodamine-conjugated antidigoxygenin antibodies. Slides were counterstained with DAPI and analyzed with an Olympus MF triple color filter.

Acknowledgments

We thank Timothy R. Mulcahy for the generous gift of the human placenta genomic λ FIX II library; and Atsushi Mizokami, Tien-Min Lin, and Pei-Wen Hsiao for their valuable discussion. We also are grateful to Loretta L. Collins for manuscript preparation. This work was supported by grants CA 55639 and DK 47258 from NIH.

References

1. Evans, R. M. (1988). *Science* **240,** 889–895.
2. Chang, C., Da Silva, S. L., Ideta, R., Lee, Y., Yeh, S., and Burbach, J. P. (1994). *Proc. Natl. Acad. Sci. USA* **91,** 6040–6044.
3. Chang, C. and Kokontis, J. (1988). *Biochem. Biophys. Res. Commun.* **155,** 971–977.
4. Chang, C., Kokontis, J., Acakpo-Satchivi, L., Liao, S., Takeda, H., and Chang, Y. (1989). *Biochem. Biophys. Res. Commun.* **165,** 735–741.
5. Kokontis, J., Liao, S., and Chang, C. (1991). *Receptor* **1,** 261–270.
6. Chang, C., Saltsman, A., Lee, H. J., Uemura, H., Sue, C., Chodak, C., et al. (1993). *Endocr. J.* **1,** 541–549.
7. Lydon, J. P., Power, R. F., and Conneely, O. M. (1992). *Gene Expression* **2,** 273–283.
8. Lee, C., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Wei, L. N. (1995). *Genomics* **30,** 46–52.
9. Becker-Andre, M., Andre, E., and DeLamarter, J. F. (1993). *Biochem. Biophys. Res. Commun.* **194,** 1371–1379.
10. Young, W. J., Smith, S. M., and Chang, C. (1997). *J. Biol. Chem.* **272,** 3109–3116.
11. Lin, T. M., Young, W. J., and Chang, C. (1995). *J. Biol. Chem.* **270,** 30,121–30,128.
12. Lee, Y. F., Pan, H. J., Burbach, J. P., Morkin, E., and Chang, C. (1997). *J. Biol. Chem.* **272,** 12,215–12,220.
13. Lee, H. J. and Chang, C. (1995). *J. Biol. Chem.* **270,** 5434–5440.
14. Lee, H. J., Lee, Y., Burbach, J. P., and Chang, C. (1995). *J. Biol. Chem.* **270,** 30,129–30,133.
15. Lee, H. J., Young, W. J., Shih, C. Y., and Chang, C. (1996). *J. Biol. Chem.* **271,** 10,405–10,412.
16. Ritchie, H. H., Wang, L. H., Tsai, S., O'Malley, B. W., and Tsai, M. J. (1990). *Nucleic Acids Res.* **18,** 6857–6862.
17. Nagata, T., Kanno, Y., Ozato, K., and Taketo, M. (1994). *Gene* **142,** 183–189.
18. Faisst, S. and Meyer, S. (1992). *Nucleic Acids Res.* **20,** 3–26.
19. Miwa, T., Boxer, L. M., and Kedes, L. (1987). *Proc. Natl. Acad. Sci. USA* **84,** 6702–6706.
20. Chodosh, L. A., Baldwin, A. S., Carthew, R. W., and Sharp, P. A. (1988). *Cell* **53,** 11–24.
21. Cortes, P., Buckbinder, L., Leza, M. A., Rak, N., Hearing, P., and Merino, A. (1988). *Genes Dev.* **2,** 975–990.
22. Smale, S. T. and Baltimore, D. (1989). *Cell* **57,** 103–113.
23. Mckusck, V. A. (1994). In: *Mendelian Inheritance in Man*, 11th ed. Johns Hopkins, Baltimore, pp. cxviii–cxix.
24. Ideta, R., Yeh, S., Lee, Y., Adachi, K., Takeda, H., Su, C., et al. (1995). *Endocrine* **3,** 277–283.
25. Kontrogianni-Konstantopoulos, A., Vlahou, A., Vu, D., and Flytzanis, C. N. (1996). *Dev. Biol.* **177,** 371–382.
26. Mitsuhashi, T. and Nikodem, V. M. (1989). *J. Biol. Chem.* **264,** 8900–8904.
27. Encio, I. J. and Detera-Wadleigh, S. D. (1991). *J. Biol. Chem.* **266,** 7182–7188.
28. Yao, T. P., Segraves, W. A., Oro, A. E., McKeown, M., and Evans, R. M. (1992). *Cell* **71,** 63–72.
29. Sharp, P. A. (1992). *Cell* **68,** 819–821.
30. Sudhof, T. C., Van der Westhuyzen, D. R., Goldstein, J. L., Brown, M. S., and Russell, D. W. (1987). *J. Biol. Chem.* **262,** 10,773–10,779.
31. Hong, F. D., Huang, H. J., To, H., Young, L. J., Oro, A., Bookstein, R., et al. (1989). *Proc. Natl. Acad. Sci. USA* **86,** 5502–5506.
32. Araki, E., Murakami, T., Shirotani, T., Kanai, F., Shinohara, Y., Shimada, F., et al. (1991). *J. Biol. Chem.* **266,** 3944–3948.
33. Pugh, B. F. and Tjian, R. (1990). *Cell* **61,** 1187–1197.
34. Lin, D. L. and Chang, C. (1996). *J. Biol. Chem.* **271,** 14,649–14,652.
35. Kern, S. E., Pietenpol, J. A., Thiagalingam, S., Seymour, A., Kinzler, K. W., and Vogelstein, B. (1992). *Science* **25,** 827–830.
36. Vogelstein, B. and Kinzler, K. W. (1992). *Cell* **70,** 523–526.
37. Murty, V. V., Houldsworth, J., Baldwin, S., Reuter, V., Hunziker, W., Besmer, P., et al. (1992). *Proc. Natl. Acad. Sci. USA* **89,** 11,006–11,010.
38. Murty, V. V., Renault, B., Falk, C. T., Bosl, G. J., Kucherlapati, R., and Chaganti, R. S. (1996). *Genomics* **35,** 562–570.
39. Tricoli, J. V., Rall, L. B., Scott, J., Bell, G. I., and Shows, T. B. (1984). *Nature* **310,** 784–786.
40. Albertson, D. G., Sherrington, P. D., and Rabbitts, P. H. (1989). *Hum. Genet.* **83,** 127–132.
41. Weiss, R. E., Weinburg, M., and Refetoff, S. (1993). **91,** 2408–2415.
42. de The, H., Marchio, A., Tiollais, P., and Dejean, A. (1987). *Nature* **330,** 667–670.
43. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). In: *Molecular Cloning*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 2.69–2.80.
44. Clark, S. S., Liang, Y., Reedstrom, C. K., and Wu, S. Q. (1994). *Blood* **84,** 4301–4309.
45. Wu, S. Q., Veolkerding, K., Sabatini, L., Chen, X. R., Huang, J. W., and Meisner, L. (1995). *Leukemia* **9,** 858–862.