Biochemical and Biophysical Research Communications **269,** 54–57 (2000)
doi:10.1006/bbrc.2000.2244, available online at http://www.idealibrary.com on IDEAL®

# Cloning and Characterization of the Human Retinoid X Receptor $\alpha$ Gene: Conservation of Structure with the Mouse Homolog

G. Li,[1] E. Walch,[1] X. Yang, S. M. Lippman, and J. L. Clifford[2]
*Department of Clinical Cancer Prevention, University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030*

Received January 14, 2000

**Retinoid X receptors (RXRs) are members of the steroid/thyroid hormone receptor superfamily which, along with retinoic acid receptors (RARs), mediate the biological effects of retinoids. These effects include the regulation of many aspects of embryonic development, reproductive and visual function, and the maintenance of epithelial homeostasis throughout life. The genes for three distinct retinoid X receptors, RXR$\alpha$, $\beta$, and $\gamma$, have been localized to separate chromosomes. In order to determine the organization of the human RXR$\alpha$ gene, we have isolated a clone containing the majority of the gene from a human genomic bacterial artificial chromosome (BAC) library and generated a physical map. The gene spans over 40 kilobases in size and contains at least 10 exons. Comparison with mapped portions of the mouse RXR$\alpha$ gene indicates highly conserved intron–exon positioning. These results provide information necessary to generate constructs for targeting the RXR$\alpha$ gene in human cell lines, which may eventually lead to an understanding of the function of RXR$\alpha$ in human cancer.** **© 2000 Academic Press**

Retinoids, including vitamin A and its metabolites, are essential for the maintenance of normal epithelial growth and differentiation, and are effective therapeutic and chemopreventive agents for several epithelial-derived cancers (1–3). Retinoids are also critically involved in embryonic development, vision, and maintenance of reproductive capacity (4, 5). Retinoids exert their biological effects primarily through two subfamilies of the steroid/thyroid hormone receptor superfamily, the retinoic acid receptors (RARs) and retinoid X receptors (RXRs) (6–8). There are three types of RARs (RAR$\alpha$, $\beta$, and $\gamma$) and RXRs (RXR$\alpha$, $\beta$, and $\gamma$), all nuclear, ligand-dependent, DNA-binding transcriptional transactivator proteins (4). Each RAR and RXR type is encoded by a different gene and has multiple splice isoforms (9). RARs are activated by all-*trans* retinoic acid (RA) and its 9-*cis* isomer (9c-RA), while RXRs are only activated by 9c-RA. Upon ligand binding, RARs and RXRs bind DNA and activate transcription primarily as RAR/RXR heterodimers (9). Ligand binding brings about conformational alterations in the receptors, resulting in the displacement of corepressor proteins and the recruitment of coactivator proteins to the RAR/RXR heterodimer (10). This results in receptor binding to DNA and transcriptional activation of a large number of genes involved in growth regulation and differentiation (4–6).

The various RARs and RXRs are highly conserved in evolution and show a greater sequence conservation for a given type across species lines than between receptor types within a species (9). The structures of several human and mouse RAR and RXR genes have been determined and the chromosomal locations for all mouse and human RARs and RXRs have been identified (6, 11–13). However, the physical structure of the human RXR$\alpha$ gene has not been reported. As part of an effort to evaluate the function of the RXR$\alpha$ protein during skin cancer progression and in normal skin physiology, we have initiated a somatic cell knockout of the RXR$\alpha$ gene in human skin-derived cell lines. The initial step toward the construction of a targeting vector required the partial mapping of the human RXR$\alpha$ gene. Here we describe the structure of the RXR$\alpha$ gene and show that it is highly conserved between mouse and human.

dbGSS/GenBank Accession Numbers: AQ917566, AQ917567, AQ917568, AQ917569, AQ917570, AQ917571, AQ917572, AQ917573, AQ917574, AQ917575, AQ917576, AQ917577, AQ917578, AQ917579, AQ917580, AQ917581, AQ917582, AQ917583.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed at University of Texas M. D. Anderson Cancer Center, Department of Clinical Cancer Prevention, Box 236, 1515 Holcombe Boulevard, Houston. TX 77030. Fax: (713)-792-0628. E-mail: jclifford@notes.mdacc.tmc.edu.

0006-291X/00 $35.00

Copyright © 2000 by Academic Press
All rights of reproduction in any form reserved.

**TABLE 1**

Primers Used for PCR and Sequencing

| Intron | 5′ primer | 3′ primer |
|---|---|---|
| 2* | CTGGGCTTCAGCACTGGCAG | ACCTTGAGGACGCCATTGAG |
| 3 | TCACCTATGAACCCCGTCAG | ATGCCCATGGCCAGGCACTT |
| 4 | CACTATGGAGTGTACAGCTG | CTCCACCTCATTCTCGTTCC |
| 5 | GGAACGAGAATGAGGTGGAG | CAGCTCTGAGAAGTGTGGGA |
| 6 | GAACGACCCTGTCACCAACA | GAGAAGGAGGCGATGAGCAGCTCATTCCAG |
| 7 | CGCCGTGAAGGACGGGATCCTCCTGGCCAC | GTCTTGTCCATCTGCATGTC |
| 8 | GAGCTTGTGTCCAAGATGCG | GACGCATAGACCTTCTCCCTC |
| 9 | GAGGGAGAAGGTCTATGCGT | AGATGTTCCAGGCATTTGAG |

*Note.* Oligonucleotide sequences were derived from the human cDNA sequence and were chosen based on the mouse gene structure to allow sequencing of the putative intron–exon boundaries.

* The size of intron 1 has not been determined.

## MATERIALS AND METHODS

*BAC library screening.* The human RXR$\alpha$ gene locus was cloned by screening a BAC (bacterial artificial chromosome) library (Genome Systems Inc., St. Louis, MO). A polymerase chain reaction (PCR)-generated radiolabeled probe containing sequences within the human RXR$\alpha$ cDNA was used to probe a human genomic BAC library. The probe was generated from the following oligonucleotides, 5′-ACTATGGAGTGTACAGCTG-3′ and 5′-CATGCCCATGGCCAGGCACTTC-3′, which were chosen based on the known intron–exon boundaries of the mouse cDNA for their likelihood to span a region within exon 4. One positive clone, pBACX$\alpha$, containing approximately 140 kb of genomic DNA was obtained and subjected to restriction mapping and sequence analysis.

*Restriction mapping.* Field Inversion Gel Electrophoresis (FIGE) was used to obtain a restriction map of pBACX$\alpha$. Purified BAC DNA was digested with infrequent cutting enzymes and FIGE gels were Southern blotted and probed with $^{32}$P-labeled oligonucleotides which were contained within each putative exon. Once the location of exons was determined within the BAC clone, subclones were subjected to fine mapping with frequent cutting enzymes.

*Long and accurate PCR (LA-PCR).* Long and accurate PCR (LA-PCR) was performed using the Takara LA-PCR kit on pBACX$\alpha$ DNA to determine the size and location of introns. Oligonucleotides were derived from the human RXR$\alpha$ cDNA sequence and designed to span putative introns based on the known intron–exon boundaries of the mouse RXR$\alpha$ gene (see oligonucleotide list in Table 1). All PCR fragments were cloned into the pCR-XL- TOPO cloning vector by the TA cloning method according to the manufacturer's instructions (Invitrogen). All clones were then sequenced at both the 5′ and 3′ ends to verify correct oligonucleotide priming and to deduce the correct intron–exon boundaries. Automated sequencing was performed with an Applied Biosystems model 373A fluorescence sequencer, at the M. D. Anderson Cancer Center Core Sequencing Facility.

## RESULTS AND DISCUSSION

The human RXR$\alpha$ gene was cloned from a lymphocyte BAC genomic library using a PCR-generated $^{32}$P-labeled probe spanning a portion of the putative exon 4, which would encode part of the DNA-binding domain. The probe was chosen based on the known intron–exon boundaries of the mouse cDNA for their likelihood to span a region within exon 4. One positive

**TABLE 2**

The Partial Nucleotide Sequences and Intron–Exon Junctions of the RXR$\alpha$ Gene

| Intron number | EXON: | Intron | Intron size (kb) | Intron | EXON |
|---|---|---|---|---|---|
| 1 | CCGCTCG | | a | tcttgcag* | ATTTCTC |
| 2 | CCCCCAG | gtgagtg | 6 | ccccc**ag** | CTCAGCT |
| 3 | TCCTCAG* | gtaccgc | 0.7 | cccgc**ag** | GCAAGCA |
| 4 | CGGGAAG | gtaggcc | 7.8 | gttgt**ag** | CCGTGCA |
| 5 | CAGCTCG | gtgagtt | 4.3 | actgc**ag** | CCGAACG |
| 6 | CGGGCAG | gtgagtg | 6.9 | tcccc**ag** | GCTGGAA |
| 7 | TTGACAG | gtggggg | 2.4 | tctgc**ag** | GGTGCTG |
| 8 | AACCCTG | gtatggc | 1.9 | ctttc**ag** | ACTCCAA |
| 9 | CGGGAAG | gtgggtc | 2.2 | cctgc**ag** | GTTCGCT |
| 10 | AGATTCC[b] | | | | |

*Note.* The intron–exon boundaries were determined by sequencing the RXR$\alpha$ BAC clone with oligonucleotides shown in Table 1. Upper and lower case letters represent exon and intron sequence, respectively. Consensus splice donor and acceptor sites are underlined. (a) Size not determined. (b) Furthest extent of the putative 3' untranslated region; 1775 base pairs downstream of the translation stop. This sequence is 105 base pairs downstream of an AAGAAA potential polyadenylation sequence. All intron and 3' untranslated region sequence has been submitted to GenBank. *Differ from previously published intron–exon boundary for mouse.

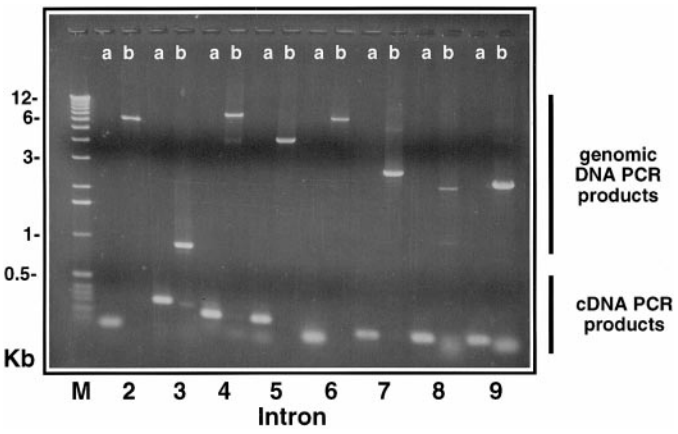


**FIG. 1.** Determination of intron size by LA-PCR. PCR products were generated using intron-spanning oligonucleotides. The a and b lanes designate PCR products generated from the RXR$\alpha$ cDNA and BAC genomic clone DNA, respectively. Numbers below correspond to introns 2–9. kb, kilobases; M, size marker.

clone, pBACX$\alpha$, containing approximately 140 kb of genomic DNA was obtained and subjected to restriction mapping and sequence analysis. After preliminary restriction mapping by FIGE (Field Inversion Gel Electrophoresis) with infrequent cutting restriction enzymes, the RXR$\alpha$ exons were localized within the BAC clone. Exon-specific primers were designed, based on the mouse intron–exon boundaries, and used to directly sequence the BAC clone (Table 1). In all but 2 cases the mouse and human intron–exon boundaries were identically located as shown in Table 2. All deduced intron–exon boundaries indicate the canonical consensus splice donor and acceptor sequences in accordance with the GT/AG rule (14).

Intron size was determined by generating PCR fragments with the LA-PCR kit (Takara) using oligonucleotide pairs situated in sequential exons (Table 1). Both BAC DNA and human genomic DNA isolated from the HaCaT cell line (15) were used as templates. The resulting fragments were subcloned into the pCR-XL-TOPO cloning vector (Invitrogen) and sequenced using vector-specific primers. Identical sized fragments were obtained for both BAC and genomic DNA templates, indicating that the BAC DNA has not been altered during cloning. PCR fragments spanning introns 2–9 (Fig. 1, b lanes and data not shown) are shown alongside fragments generated by the same oligonucleotides using the hRXR$\alpha$ cDNA as template (Fig. 1, a lanes). All cDNA template fragment sizes are as expected based on the known sequence (6). A PCR fragment spanning intron sequence between exons 1 and 2 could not be generated from either BAC or genomic DNA. Based on the structure of the mouse RXR$\alpha$ gene (11) we speculate that there is a large distance between these exons (>40 kb) which would place exon 1 outside of the BAC clone. This is likely since exon 1-specific oligonucleotides do not detect bands on southern blots of BAC DNA (data not shown). Alternatively, if exon 1 is contained within the BAC clone, the distance between exons 1 and 2 may still exceed the length possible for efficient PCR extension.

The human RXR$\alpha$ gene consists of at least 10 exons separated by introns ranging in size from 700 base pairs (intron 3) to more than 7.8 kb (intron 4) (Fig. 2B). Comparison of the human gene to the previously published portion of the mouse gene (16) indicates a nearly identical structure (Figs. 2B and 2C). This is consistent with the conserved intron–exon boundaries (Table 2) and the extremely high degree of amino acid conservation between the human and mouse RXR$\alpha$ proteins (9). A close similarity in gene organization between the mouse and human RXR$\beta$ gene was also observed by Nagata *et al.* (12). The position of three putative 5′ exons (1a-c) is indicated in Fig. 2B. The mouse RXR$\alpha$ gene, while not completely mapped, has been shown to encode at least three alternate 5′ exons which can generate splice variant mRNAs, mRXR$\alpha$1, $\alpha$2 and $\alpha$3, the latter two of which have only been detected in the testes (13). The organization of these exons in this region of the mouse RXR$\alpha$ gene has not been delineated. We have attempted without success to detect similar alternative 5′ exons in the human gene by probing BAC DNA Southern blots with oligonucleotides derived from the mouse alternate 5′ exon sequences. Also, attempts to generate PCR fragments using these same oligonucleotides in combination with exon 2-specific oligonucleotides have failed. It is likely that such isoforms exist in humans and further analysis of the human RXR$\alpha$ gene with additional BAC genomic clones will be necessary in order to fully map the 5′ region.

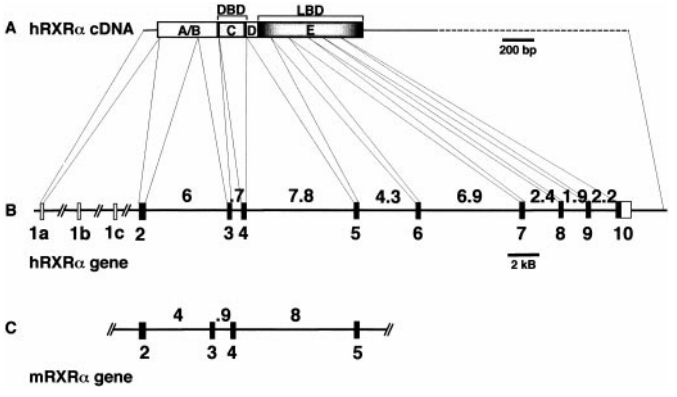


**FIG. 2.** Physical map of the human RXR$\alpha$ gene. (A) The human RXR$\alpha$ cDNA sequence, including putative 3′ untranslated sequence (dotted line). (B) The human RXR$\alpha$ gene structure. Numbered boxes represent exons and the lines represent introns. Exons 1b and 1c are putative, based on the known mouse alternate splice mRNAs. Numbers above the line represent intron sizes in kb. (C) The known mouse RXR$\alpha$ gene locus. Cross hatches represent regions of unknown size.

Sequence analysis of exon 10 indicates a long 3′-untranslated region which does not contain the canonical AATAAA polyadenylation signal. However a potential noncanonical AAGAAA sequence was detected 1676 base pairs downstream of the translation stop codon (Fig. 2B and Table 2). No polyadenine tail was detected as far as 105 base pairs downstream of the AAGAAA sequence, allowing for the likelihood that the 3′-untranslated region is longer than shown.

In summary we have determined the partial structure of the human RXR$\alpha$ gene as part of an effort to generate a targeting construct for somatic cell knockouts in human cell lines. A similar somatic cell knockout approach, aimed at understanding RXR$\alpha$ function in mouse embryonal carcinoma cells, has been quite fruitful (8, 17). These same cells have also proven useful for determining receptor-dependent and -independent effects of retinamides, a class of retinoids with potential for cancer treatment and prevention (18). However, differences in the physiology of human and mouse tissues make a direct comparison of these results to the human situation potentially inaccurate. For example, the role of retinoid signaling in the progression of nonmelanoma skin cancer in humans may differ from that of mice, which do not get this type of cancer under normal environmental conditions. The data gathered from the study of retinoid receptor knockout mice in this context might not reflect every aspect of retinoid receptor function in human skin cancer. The physical structure of the human RXR$\alpha$ gene determined in this study will allow the development of targeting vectors for the production of human somatic cell knockouts. Analysis of these cells will yield novel information concerning the function of RXR$\alpha$ in human cancer.

## ACKNOWLEDGMENTS

The authors thank Reuben Lotan and Vemparala Subbarayan for helpful discussions and David Menter for critical reading of the manuscript. This work was supported in part by NCI CA78560-01.

## REFERENCES

1. Smith, M. A., Parkinson, D. R., Cheson, B. D., and Friedman, M. A. (1992) *J. Clin. Oncol.* **10,** 839–864.
2. Lotan, R. (1996) *FASEB J.* **10,** 1031–1039.
3. Lippman, S. M., and Davies, P. J. (1997) *Cancer Chemother. Biol. Response Modifiers* **17,** 349–362.
4. Chambon, P. (1996) *FASEB J.* **10,** 940–954.
5. Gudas, L. J., Sporn, M. B., and Roberts, A. B. (1994) *in* The Retinoids: Biology, Chemistry and Medicine (Sporn, M. B., Roberts, A. B., and Goodman, D. S., Eds.), pp. 443–520, Raven Press, Ltd., New York, NY.
6. Mangelsdorf, D. J., Umesono, K., and Evans, R. M. (1994) *in* The Retinoids: Biology, Chemistry and Medicine (Sporn, M. B., Roberts, A. B., and Goodman, D. S., Eds.), pp. 443–520, Raven Press, Ltd., New York, NY.
7. Kastner, P., Mark, M., and Chambon, P. (1995) *Cell* **83,** 859–869.
8. Clifford, J. L., and Lippman, S. M. (1997) *J. Oncol. Index Rev.* **1,** 3–5.
9. Leid, M., Kastner, P., and Chambon, P. (1992) *Trends Biochem. Sci.* **17,** 427–433.
10. Xu, L., Glass, C. K., and Rosenfeld, M. G. (1999) *Curr. Opin. Genet. Dev.* **9,** 140–147.
11. Liu, Q., and Linney, E. (1993) *Mol. Endocrinol.* **7,** 651–658.
12. Nagata, T., Weiss, E. H., Abe, K., Kitagawa, K., Ando, A., Yara-Kikuti, Y., Seldin, M. F., Ozato, K., Inoko, H., and Taketo, M. (1995) *Immunogenetics* **41,** 83–90.
13. Brocard, J., Kastner, P., and Chambon, P. (1996) *Biochem. Biophys. Res. Commun.* **229,** 211–218.
14. Shapiro, M. B., and Senapathy, P. (1987) *Nucleic Acids Res.* **15,** 7155–7174.
15. Boukamp, P., Petrussevska, R. T., Breitkreutz, D., Hornung, J., Markham, A., and Fusenig, N. E. (1988) *J. Cell Biol.* **106,** 761–771.
16. Kastner, P., Grondona, J. M., Mark, M., Gansmuller, A., LeMeur, M., Decimo, D., Vonesch, J. L., Dolle, P., and Chambon, P. (1994) *Cell* **78,** 987–1003.
17. Clifford, J., Chiba, H., Sobieszczuk, D., Metzger, D., and Chambon, P. (1996) *EMBO J.* **15,** 4142–4155.
18. Clifford, J. L., Menter, D. G., Wang, M., Lotan, R., and Lippman, S. M. (1999) *Cancer Res.* **59,** 14–18.