Vol. 171, No. 2, 1990 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
September 14, 1990 Pages 697-704

MOLECULAR CLONING OF ANDROGEN RECEPTORS FROM DIVERGENT SPECIES WITH A POLYMERASE CHAIN REACTION TECHNIQUE: COMPLETE cDNA SEQUENCE OF THE MOUSE ANDROGEN RECEPTOR AND ISOLATION OF ANDROGEN RECEPTER cDNA PROBES FROM DOG, GUINEA PIG AND CLAWED FROG+

Wei Wu He[1,2], Leslie M. Fischer[3], Siquan Sun[1], David L. Bilhartz[1], Xiaoping Zhu[1], Charles Y-F. Young[1], Darcy B. Kelley[3], and Donald J. Tindall[1]*

[1]Department of Urology/Biochemistry and Molecular Biology, Mayo Foundation, Rochester, MN 55905

[2]Department of Cell Biology, Baylor College of Medicine, Houston, TX 77030

[3]Department of Biological Sciences, Columbia University, New York, NY 10027

Received August 3, 1990

Summary: We have cloned and sequenced 2.8 kilobases of cDNA encoding the mouse androgen receptor by RNA amplification with transcript sequencing. Sequence analysis predicts that this cDNA contains an open reading frame of 2697 nucleotides encoding a polypeptide of 899 amino acids. Androgen receptor cDNA probes of dog, guinea pig, and frog were also isolated and sequenced using consensus primers derived from human and rat androgen receptor cDNA s. Northern blot analysis with the species-specific probes revealed similarities in size between amphibian and mammalian mRNA s. These results demonstrate the utility of this technique in obtaining nucleic acid probes and sequence information of steroid receptors from different species. The sequence data and the Northern blot analysis of the receptors in different species demonstrate that the androgen receptor has been well-conserved during evolution.

Androgen action has been studied widely in different species due to its important role in male physiology and pathology. Species-specific molecular probes for AR are required to address biological questions in a particular species at the molecular level. Since cDNA probes for both the human and rat AR are available (1,2,3,4,5,6) molecular probes for AR in different species can be obtained by conventional cloning methods. However, isolation of cDNA clones for a rare mRNA, such as that of AR, has proven to be both time consuming and tedious. The general use of PCR technology to amplify a particular DNA sequence from different species has been

+ The accession number for the mouse AR sequence in the EMBL/Gen Bank Databases is X53779.

* To whom correspondence and reprint requests should be addressed.

Abbreviations: (m)AR: (mouse) androgen receptor; RAWTS: RNA amplification with transcript sequencing; kb: kilobases; PCR: polymerase chain reaction.

0006-291X/90 $1.50

suggested (7). Recently, the utility of the PCR technique in analyzing mitochondrial and other DNA sequences in different species has been reported (8). In this report, the complete mouse AR coding region cDNA and partial cDNA probes for dog, guinea pig, and frog have been isolated and sequenced directly with RAWTS (9), which involves PCR amplification of cDNA with primers containing a phage promoter and generation of RNA from the promoter for direct sequencing. By using this technique, we have demonstrated that both DNA and RNA probes, together with nucleic acid sequence information for AR, can be obtained from different species without conventional cloning. Thus, species-specific probes can be used to address biological questions in different species, and the sequence information can facilitate the analysis of evolutionary changes of functional domains of the receptor.

MATERIALS AND METHODS

Restriction enzymes, avian myeloblastosis virus reverse transcriptase, klenow DNA polymerase, T_7 RNA polymerase, T_4 DNA ligase, plasmid pGEM-4Z, pGEM-11Zf(+), Erase-a-base system, Riboprobe Gemini System, TaqTrack Sequencing Systems and a λgt11 mouse kidney cDNA library were obtained from Promega Corporation. [α^{35}S] dATP (400 Ci/mmol), [α^{32}P] dCTP (3000 Ci/mmol), [γ^{32}P] dATP (5000Ci/mmol), Multiprime DNA Labeling Systems were obtained from Amersham International. Mouse EMBL3 genomic library, mouse testis poly(A) RNA, dog kidney poly(A) RNA, guinea pig testis poly(A) RNA were obtained from Clontech Laboratories, Inc. Reagents for PCR were obtained from Perkin Elmer Cetus. Gene Clean kits were obtained from Bio101. Zeta-Probe membranes for Southern and Northern blot analysis were obtained from Bio-Rad. All other chemicals were obtained from Sigma chemical company. Oligonucleotide primers were synthesized with the Applied Biosystems 380A DNA synthesizer. The sequences of primers used in this report are: mAR (1-15) U (5'GAATTCGGTGGAAGCT3'), mAR (943-959) T7 D (5'TAATACGACTCACTATAGGGAG AGTAACCTCCCTTGAAAG3') for the amplification of fragment mARpa, mAR (787-802) U (5'AA CATCTGAGTCCAGG3'), mAR (1577-1592) T7 D (5'TAATACGACTCACTATAGGGA GACATGTCCCCATAAGGT3') for mARpb, mAR (1409-1424) U (5'TATGGCTACACTCG GC3'), mAR (2761-2777) T7 D (5'TAATACGACTCACTATAGGGAGAAAAGGGAACAAGG T3') for mARpc (Figure 1), UniAR U SP6 (5'GATTTA GGTGACACTATAGAACTGCAGTGT GGCAGCTGCAAAGTCTT3'), UNIAR D T7 (5'T AATACGACTCACTATAGGGAGACTCG AGTTCCTTGATGTAGTTCAT3') for the amplification of UniAR fragment (Figure 3). The numbers used in the names of the primers represent the position of primer on the mAR cDNA sequence. T_7 (or SP_6) represents the primer which contains T_7 (or Sp_6) promotor sequence. D (or U) represents the primer which is downstream (or upstream) for PCR amplification. Internal sequencing primers are 15-20 oligonucleotides derived from the mouse AR cDNA sequence.

Synthesis of cDNA, amplification of the cDNA with PCR and direct sequence analysis of the PCR amplified product was carried out as described (9). A Balb/c mouse kidney λgt11 cDNA library was screened with a HindIII and EcoRI rat AR cDNA probe containing part of the DNA and steroid binding domains with standard techniques (10). The probe was labeled with the Multiprime DNA Labelling System at a specific activity of > $1x10^9$ dpm/μg. Hybridization was carried out with 5X TEN (20X TEN: 0.3M Tris-HCl pH8.0, 0.02M EDTA and 3M NaCl), 5X Denhardts, 0.5% Sodium pyrophosphate, 0.1% SDS, 0.2mg/ml heat denatured salmon sperm DNA and $1X10^6$ cpm/ml of [^{32}P]-labeled rat AR probe at 55°C for 12 hours. The filters were washed in 0.5X TEN at room temperature for 20-30 min, then at 55°C for 15 min, and finally at room temperature for 10 min. The filters were dried and autoradiographed at -70°C using Kodak XAR-5 film with two intensifying screens. The positive clones were plaque-purified, and the inserts were subcloned into pGEM-11Zf(+). The genomic clones were obtained in the same manner as described above, except that the EMBL3 mouse genomic library from Clontech was screened with a human AR N-terminus EcoRI-BsmI fragment (Figure 1). The positive genomic clones were mapped with EcoRI and BamHI. A 1.7 kb EcoR1 fragment recognized by the human androgen receptor N-terminus probe on a southern blot was subcloned into pGEM-4Z.

The 1.7 kb genomic fragment in pGEM-4Z and the 1.2 kb cDNA fragment in pGEM-11Zf(+) were sequenced with the double strand dideoxy sequencing method with Sequenase (11). Nested deletions were constructed to sequence the 1.9 kb fragment with the Erase-a-base System from Promega according to the manufacturer's protocol. Ambiguous regions observed with reverse transcriptase and Sequenase were sequenced with Taq polymerase with the TaqTrack system.

Southern blot analysis of PCR products was carried out with approximately 300ng of Gene clean purified UniAR DNA, amplified from cDNA of different species. After electrophoresis on a 1.5% agarose gel, Southern blot analysis was carried out as described (12) with minor modifications. Briefly, the nylon filter blots were hybridized in 0.5 M $NaHPO_4$, pH7.2, 1 mM EDTA, 7% sodium dodecyl sulfate and $1X10^6$ cpm/ml [^{32}P]-labeled mAR Hinc II-EcoR I fragment at 55°C for 12 hours. Filters were washed in 40 mM $NaHPO_4$, 1 mM EDTA and 1% SDS at 55°C five times for 30 minutes each before autoradiography at -70°C with Kodak XAR-5 film for 12 hours.

Total RNA was isolated from human LNCaP cells, adult male South African clawed frog lung, liver, kidney, testis, thigh and juvenile larynx by the method of Chirgwin (13). The transfer of RNA and hybridization were carried out as described (10).

RESULTS

Amplification, Cloning and Sequencing of the Mouse Androgen Receptor cDNA. The strategy to amplify and directly sequence the mAR full-length cDNA by the RAWTS technique is illustrated in Figure 1. The putative mAR cDNA was divided into three segments for the convenience of PCR amplification. For each segment two oligonucleotide primers were designed from homologous regions of the human and rat AR's. The downstream primer of each pair contained a phage T_7 promotor sequence. Total RNA was isolated from the kidneys of Balb/c mice and used to synthesize a first strand cDNA. Three anticipated segments (mARpa, mARpb and mARpc in Figure 1) were amplified with the primers. These fragments were purified, transcribed into RNA with T_7 RNA polymerase and sequenced directly with the upstream primer and internal primers using reverse transcriptase (9). Regions which were difficult to read by this method were sequenced with Taq polymerase.

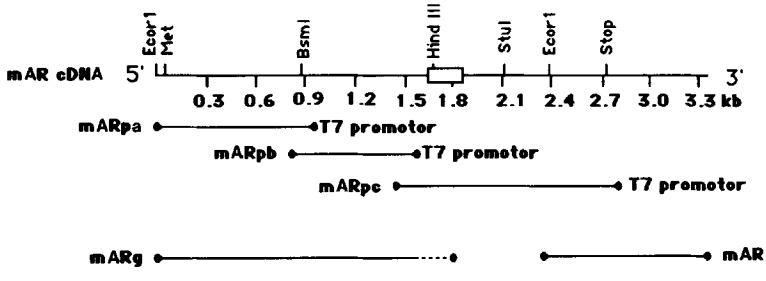


Fig.1. Restriction Map of the Mouse Androgen Receptor cDNA and Genomic DNA. mARpa, mARpb and mARpc denote PCR products amplified from mouse kidney cDNA with primers derived from rat and human sequences. mARg denotes an EcoR1 fragment of a mouse androgen receptor genomic clone isolated from a mouse EMBL 3 library. mARg contains the first exon-intron boundary and the dotted line stands for intron sequences. mAR1 denotes a cDNA clone isolated from a lambda gt-11 mouse kidney cDNA library. The predicted initiation methionine (Met), the in-frame stop codon (stop), and the DNA-binding domain ⊏⊐ are indicated.

In order to test the fidelity of the DNA sequences obtained by the PCR technique, both genomic and cDNA clones were isolated and sequenced. A 1.7 kb EcoR1 fragment (mARg, Figure 1) from one of the clones isolated from a Balb/c mouse EMBL3 genomic library and a 1.2 kb cDNA clone (mAR1, Figure 1) isolated from a mouse kidney cDNA library were subcloned and sequenced with the dideoxy termination method by constructing nested-deletions of the inserts. The sequences obtained by conventional cloning techniques were found to be identical to those obtained by direct sequencing of the PCR amplified mouse AR cDNA products.

The complete nucleic acid sequence of the mAR obtained by both PCR techniques and conventional cloning techniques is shown in Figure 2. The mouse sequence contains a predicted translation initiation codon at nucleotide 33, which would initiate an open reading frame of 899 amino acids. The predicted molecular weight of the mAR protein is 98,000. The mouse AR has 97% and 86% overall amino acid sequence similarities with the rat and human AR, respectively. The putative DNA and steroid binding domains of the mAR are identical to that of the human and rat AR.

Amplification and Sequence Comparison of Androgen Receptor cDNA in other Species. In order to obtain probes to study androgen action in more divergent species, RAWTS was used to amplify and sequence directly several AR cDNA probes. Two oligonucleotide primers were designed from regions of the human, rat and mouse AR which had high sequence similarities. The upstream primer was derived from the first zinc-finger region, which was predicted to be well-conserved among all the species. The downstream primer was derived from a region in the steroid binding domain, where the corresponding region in the estrogen receptor had been shown to be conserved in different species (14). The upstream and downstream primers were linked to bacterial SP_6 and T_7 promoters, respectively; so that the amplified products would contain both SP_6 and T_7 promoters at either end. These two consensus primers (UniAR primers) were used to amplify cDNA from dog, guinea pig, and frog. A predicted band of 800 nucleotides was present in all the species amplified. Southern blot analysis revealed that each of these cDNA's was recognized by the mouse androgen receptor probe (Figure 3). However, the intensity of the signal on the frog DNA was significantly less than the other species, even though the same amount of DNA was loaded on the agarose gel. This suggests that the frog AR has less sequence similarity with the mouse AR than either the dog or the guinea pig AR's.

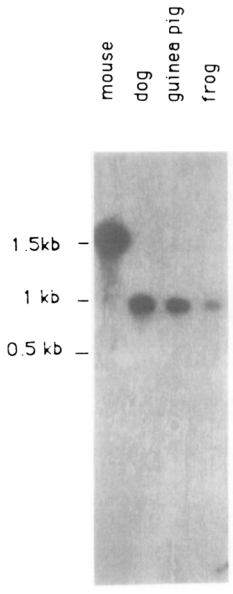
The amplified products of the AR fragments from different species were transcribed into RNA with T_7 RNA polymerase, and the transcribed RNA's were used as templates for direct sequencing with an SP_6 promotor as primer. The sequences obtained from dog, guinea pig, and frog were compared with that of human, rat and mouse AR (Figure 4). At the DNA level, the sequence similarities with the human are the following: rat (91%), mouse (95%), dog (95%), guinea pig (93%) and frog (86%). However, the deduced amino acid sequences of this region were found to be identical in all six species, with the exception of the third amino acid, glycine, which is an alanine in the frog. Most of the nucleotide changes were found in the third nucleotide base and did not change the amino acid.

```
                                                                                                    GAATTCGGTGGAAGCTACAGACAAGCTCAAGG   32

    ATGGAGGTGCAGTTAGGGCTGGGAAGGGTCTACCCACGGCCCCCATCCAAGACCTATCGAGGAGCGTTCCAGAATCTGTTCCAGAGCGTGCGCGAAGCGATCCAGAACCCGGGCCCCAGG  152
1   MetGluValGlnLeuGlyLeuGlyArgValTyrProArgProProSerLysThrTyrArgGlyAlaPheGlnAsnLeuPheGlnSerValArgGluAlaIleGlnAsnProGlyProArg

    CACCCTGAGGCCCGCTAACATAGCACCTCCCGGCGCCTGTTTACAGCAGAGGCAGGAGACTAGCCCCCGGCGGCGGCGGCGGCAGCAGCACACTGAGGATGGTTCTCCTCAAGCCCACATC  272
41  HisProGluAlaAlaAsnIleAlaProProGlyAlaCysLeuGlnGlnArgGlnGluThrSerProArgArgArgArgArgGlnGlnHisThrGluAspGlySerProGlnAlaHisIle

    AGAGGCCCCACAGGCTACCTGGCCCTGGAGGAGGAACAGCAGCCTTCACAGCAGCAGGCAGCCTCCGAGGGCCACCCTGAGAGCAGCTGCCTCCCCGAGCCTGGGGCGGCCACCGCTCCT  392
81  ArgGlyProThrGlyTyrLeuAlaLeuGluGluGluGlnGlnProSerGlnGlnGlnAlaAlaSerGluGlyHisProGluSerSerCysLeuProGluProGlyAlaAlaThrAlaPro

    GGCAAGGGGCTGCCGCAGCAGCCACCAGCTCCTCCAGATCAGGATGACTCAGCTGCCCCATCCACGTTGTCCCTGCTGGGCCCCACTTTCCCAGGCTTAAGCAGCTGCTCCGCCGACATT  512
121 GlyLysGlyLeuProGlnGlnProProAlaProProAspGlnAspAspSerAlaAlaProSerThrLeuSerLeuLeuGlyProThrPheProGlyLeuSerSerCysSerAlaAspIle

    AAAGACATTTTGAACGAGGCCGGCACCATGCAACTTCTTCAGCAGCAGCAACAACAGCAGCAGCACCAACAGCAGCACCAACAGCACCAACAGCAGCAGGAGGTAATCTCCGAAGGCAGC  632
161 LysAspIleLeuAsnGluAlaGlyThrMetGlnLeuLeuGlnGlnGlnGlnGlnGlnHisGlnGlnGlnHisGlnGlnGlnGlnGluValIleSerGluGlySer

    AGCGCAAGAGCCAGCGAGGCCACGGGGGCTCCCTCTTCCTCCAAGGATAGTTACCTAGGGGGCAATTCAACCATATCTGACAGTGCCAAGGAGTTGTGTAAAGCAGTGTCTGTGTCCATG  752
201 SerAlaArgAlaArgGluAlaThrGlyAlaProSerSerSerLysAspSerTyrLeuGlyGlyAsnSerThrIleSerAspSerAlaLysGluLeuCysLysAlaValSerValSerMet

    GGATTGGGTGTGGAAGCATTGGAACATCTGAGTCCAGGGGAACAGCTTCGGGGAGACTGCATGTACGCGTCGCTCCTGGGAGGTCCACCCGCGGTGCGTCCCACTCCTTGTGCGCCGCTG  872
241 GlyLeuGlyValGluAlaLeuGluHisLeuSerProGlyGluGlnLeuArgGlyAspCysMetTyrAlaSerLeuLeuGlyGlyProProAlaValArgProThrProCysAlaProLeu

    CCCGAATGCAAAGGTCTTCCCCTGGACGAAGGCCCAGGCAAAAGCACTGAAGAGACTGCTGAGTATTCCTCTTTCAAGGGAGGTTACGCCAAAGGATTGGAAGGTGAGAGCTTGGGGTGC  992
281 ProGluCysLysGlyLeuProLeuAspGluGlyProGlyLysSerThrGluGluThrAlaGluTyrSerSerPheLysGlyGlyTyrAlaLysGlyLeuGluGlySerLeuGlyCys

    TCTGGCAGCAGTGAAGCAGGTAGCTCTGGGACACTTGAGATCCCGTCCTCTCTGTCTCTGTATAAATCTGGAGCACTAGACGAGGCAGCAGCATACCAGAATCGCGACTACTACAACTTT 1112
321 SerGlySerSerGluAlaGlySerSerGlyThrLeuGluIleProSerSerLeuSerLeuTyrLysSerGlyAlaLeuAspGluAlaAlaAlaTyrGlnAsnArgAspTyrTyrAsnPhe

    CCGCTGGCTCTGTCCGGGCCGCCGCACCCCCCGCCCCCTACCCATCCACACGCCCGTATCAAGCTGGAGAACCCATTGGACTACGGCAGCGCCTGGGCTGCGGCGGCAGCGCAATGCCGC 1232
361 ProLeuAlaLeuSerGlyProProHisProProProProThrHisProHisAlaArgIleLysLeuGluAsnProLeuAspTyrGlySerAlaTrpAlaAlaAlaAlaAlaGlnCysArg

    TATGGGGACTTGGGTAGTCTACATGGAGGGAGTGTAGCCGGGCCCAGCACTGGATCGCCCCCAGCCACCACCTCTTCTTCCTGGCATACTCTCTTCACAGCTGAAGAAGGCCAATTATAT 1352
401 TyrGlyAspLeuGlySerLeuHisGlyGlySerValAlaGlyProSerThrGlySerProProAlaThrThrSerSerSerTrpHisThrLeuPheThrAlaGluGluGlyGlnLeuTyr

    GGGCCAGGAGGCGGGGGCGGCAGCAGCCCAAGCGATGCCGGGCCTGTAGCCCCCTATGGCTACACTCGGCCCCCTCAGGGGCTGACAAGCCAGGAGAGTGACTACTCTGCCTCCGAA 1472
441 GlyProGlyGlyGlyGlyGlySerSerSerProSerAspAlaGlyProValAlaProTyrGlyTyrThrArgProProGlnGlyLeuThrSerGlnGluSerAspTyrSerAlaSerGlu

    GTGTGGTATCCTGGTGGAGTTGTGAACAGAGTACCCTATCCCAGTCCCAATTGTGTCAAAAGTGAAATGGGACCTTGGATGGAGAACTACTCCGGACCTTATGGGGACATGCGTTTGGAC 1592
481 ValTrpTyrProGlyGlyValValAsnArgValProTyrProSerProAsnCysValLysSerGluMetGlyProTrpMetGluAsnTyrSerGlyProTyrGlyAspMetArgLeuAsp

    AGTACCAGGGACCATGTTTTACCCATCGACTATTACTTTCCACCCCAGAAGACCTGCCTGATCTGTGGAGATGAAGCTTCTGGCTGTCACTACGGAGCTCTCACTTGTGGCAGCTGCAAG 1712
521 SerThrArgAspHisValLeuProIleAspTyrTyrPheProProGlnLysThrCysLeuIleCysGlyAspGluAlaSerGlyCysHisTyrGlyAlaLeuThrCysGlySerCysLys

    GTCTTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTATCTATGTGCCAGCAGAAACGATTGTACCATTGATAAATTTCGGAGGAAAAATTGCCCATCTTGTCGTCTCCGGAAATGTTAT 1832
561 ValPhePheLysArgAlaAlaGluGlyLysGlnLysTyrLeuCysAlaSerArgAsnAspCysThrIleAspLysPheArgArgLysAsnCysProSerCysArgLeuArgLysCysTyr

    GAAGCAGGGATGACTCTGGGAGCTCGTAAGCTGAAGAAACTTGGAAATCTAAAACTACAGGAGGAAGGAGAAAACTCCAATGCTGGCAGCCCCACTGAGGACCCATCCCAGAAGATGACT 1952
601 GluAlaGlyMetThrLeuGlyAlaArgLysLeuLysLeuGlyAsnLeuLysLeuGlnGluGluGlyGluAsnSerAsnAlaGlySerProThrGluAspProSerGlnLysMetThr

    GTATCACACATTGAAGGCTATGAATGTCAGCCTATCTTTCTTAACGTCCTGGAAGCCATTGAGCCAGGAGTGGTGTGTGCCGGACATGACAACAACCAACCAGATTCCTTTGCTGCCTTG 2072
641 ValSerHisIleGluGlyTyrGluCysGlnProIlePheLeuAsnValLeuGluAlaIleGluProGlyValValCysAlaGlyHisAspAsnAsnGlnProAspSerPheAlaAlaLeu

    TTATCTAGCCTCAATGAGCTTGGAGAGAGGCAGCTTGTGCATGTGGTCAAGTGGGCCAAGGCCTTGCCTGGCTTCCGCAACTTGCATGTGGATGACCAGATGGCGGTCATTCAGTATTCC 2192
681 LeuSerSerLeuAsnGluLeuGlyGluArgGlnLeuValHisValValLysTrpAlaLysAlaLeuProGlyPheArgAsnLeuHisValAspAspGlnMetAlaValIleGlnTyrSer

    TGGATGGGACTGATGGTATTTGCCATGGGTTGGCGGTCCTTCACTAATGTCAACTCCAGGATGCTCTACTTTGCACCTGACTTGGTTTTCAATGAGTACCGCATGCACAAGTCTCGGATG 2312
721 TrpMetGlyLeuMetValPheAlaMetGlyTrpArgSerPheThrAsnValAsnSerArgMetLeuTyrPheAlaProAspLeuValPheAsnGluTyrArgMetHisLysSerArgMet

    TACAGCCAGTGTGTGAGGATGAGGCACCTGTCTCAAGAGTTTGGATGGCTCCAAATAACCCCCCAGGAATTCCTGTGCATGAAAGCACTGCTGCTCTTCAGCATTATTCCAGTGGATGGG 2432
761 TyrSerGlnCysValArgMetArgHisLeuSerGlnGluPheGlyTrpLeuGlnIleThrProGlnGluPheLeuCysMetLysAlaLeuLeuLeuPheSerIleIleProValAspGly

    CTGAAAAATCAAAAATTCTTTGATGAACTTCGAATGAACTACATCAAGGAACTCGATCGCATCATTGCATGCAAAAGAAAGAATCCCACATCCTGCTCAAGGCGCTTCTACCAGCTCACC 2552
801 LeuLysAsnGlnLysPhePheAspGluLeuArgMetAsnTyrIleLysGluLeuAspArgIleIleAlaCysLysArgLysAsnProThrSerCysSerArgArgPheTyrGlnLeuThr

    AAGCTCCTGGATTCTGTGCAGCCTATTGCAAGAGAGCTGCATCAGTTCACTTTTGACCTGCTAATCAAGTCCCATATGGTGAGCGTGGACTTTCCTGAAATGATGGCAGAGATCATCTCT 2672
841 LysLeuLeuAspSerValGlnProIleAlaArgGluLeuHisGlnPheThrPheAspLeuLeuIleLysSerHisMetValSerValAspPheProGluMetMetAlaGluIleIleSer

    GTGCAAGTGCCCAAGATCCTTTCTGGGAAAGTCAAGCCCATCTATTTCCACACACAGTGAAGATTTGGAAACCCTAATACCCAAAACCCACCTTGTTCCCTTTCCAGATGTCTTCTGCCT 2792
881 ValGlnValProLysIleLeuSerGlyLysValLysProIleTyrPheHisThrGlnEnd

    GTTATATAACTCTGCACTACTTCTCTGCAGTGCCTTGGGGGAAATTCCTCTACTGATGTACAGTCTGTCGTGAACAGGTTCCTCAGTTCTATTTCCTGGGCTTCTCCTTCTTTTTTTTTC 2912
    TTCTTCCCTCCCTCTTTCACCCTCCCATGGCACATTTTGAATCTGCTGCTGATTGTGGCTCTGCCTTTGTTTTGATTTCTGTTGTATT
```

Fig.2. Nucleotide Sequence Analysis and the Deduced Amino Acid Sequence of the cDNA Encoding the Mouse Androgen Receptor.

The numbers in the right-hand column refer to the nucleotide sequence, and the numbers in the left-hand column refer to the amino acid sequence.

Northern Blot Analysis of the Androgen Receptors in Different Species. The [^{32}P]-labeled 0.8-kb frog AR fragment was used as a hybridization probe to examine the tissue distribution of AR mRNA in the frog. A 9.5-kb mRNA was found as a predominant band in the total RNA isolated from frog larynx, liver, kidney and testis (Figure 5). The size of this mRNA is similar to that shown previously for both the human and rat AR (2,15). Low levels of AR mRNA were detected



③

```
                547
                AlaSerGlyCysHisTyrGlyAlaLeuThrCysGlySerCysLysVa
Human           GCTTCTGGGTGTCACTATGGAGCTCTCACATGTGGAAGCTGCAAGGT
Rat             ........T........C...........T.....C...........
Mouse           ........C........C...........T.....C...........
Dog             ........C....................T..............A..
G.Pig           ..A.....................C.....T.....C...........
Frog            ..C...C..C...........C..T..G.....C........A..
                      Ala

                563
                PhePheLysArgAlaAlaGluGlyLysGlnLysTyrLeuCysAlaSe
Human           TTCTTCAAAAGAGCCGCTGAAGGGAAACAGAAGTACCTGTGCGCCAG
Rat             ..............T..G.................T..A..T.....
Mouse           ....................................T.....
Dog             .....T..............................T.....
G.Pig           ..............T........A.............T.....
Frog            ........G.....A.................A..T..T........
```

④

Fig.3. Amplification of Androgen Receptor cDNA from Different Species.

Southern blot analysis of the PCR products amplified from dog, guinea pig and frog cDNA with the UniAR primers. The lane marked "mouse" contains purified mARpc products. The lanes marked "dog", "guinea pig" and "frog" contain purified UniAR products amplified from different species. A Stu I-EcoR I fragment of the mouse androgen receptor cDNA was used as probe.

Fig.4. Nucleic Acid Sequences from Part of the DNA-binding Domain of the Androgen Receptors from Different Species.

The dog, guinea pig and frog sequences were obtained by sequencing the PCR products directly. The number refers to the amino acid sequence position of the mouse androgen receptor. The last two cys residues in the first zinc finger are indicated by bold lettering.

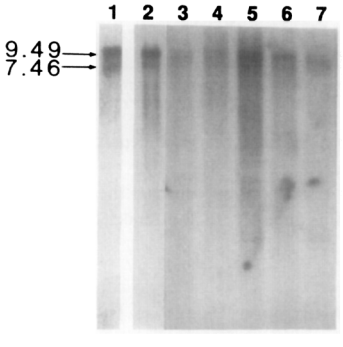


Fig.5. Northern Analysis of Human and Frog Androgen Receptor RNA.

Twenty μg of total RNA from the human LNCap cell line (lane 1), and 10 μg of total RNA from male frog larynx, lung, liver, kidney, testis, thigh (lanes 2-7, respectively) were electrophoresed, transferred to nylon membrane, and hybridized with a human androgen receptor cDNA probe for the human RNA and the frog uniAR probe for the frog RNA. The RNA molecular markers 9.49 and 7.46 Kb are shown on the left-hand side.

in the total RNA from frog lung and thigh. Of all the tissues examined in the frog, AR mRNA levels were highest in the larynx. In Figure 5, lanes 2-7 were hybridized on one membrane; however, lane 2 was exposed for only 2 days while lanes 3-7 were exposed for 15 days. A smaller sized band of about 8.0-kb was also observed in frog larynx under high hybridization stringency conditions. A similar 8.0-kb band was detected with a [^{32}P]-labeled human AR Hind III-EcoRI fragment in the human LNCaP cell line, which also contains higher amounts of AR than normal tissues (Figure 5, Lane 1). We also examined AR mRNA expression in mouse and rat tissues. A major band of 9.5 kb was observed in all the species examined, and the results agreed with those published previously (2,5,15).

DISCUSSION

By using consensus primers for the AR, species-specific DNA probes were obtained easily by amplification with the PCR technique. This approach, which has been used to amplify mitochondrial DNA from different species (8), has great potential for steroid receptor studies. Since steroid action is widely studied in different species, this approach will enable one to obtain a species-specific probe in order to address unique biological questions in a particular animal model. For example, certain muscles such as the frog larynx and specific groups of neurons in the frog brain have been shown to be target areas for androgens and to participate in determining differential sexual behavior in the frog (18,19,20). The regulation of AR in these androgen target areas is poorly studied due to the lack of appropriate molecular probes. Using the approach described in this report, frog AR DNA and RNA probes can be obtained using consensus primers and PCR amplification. The PCR products containing SP_6 and the T_7 promoter sequence serve most of the functions of a plasmid, such as the pGEM plasmid. RNA probes can be made easily from the PCR products for labeling and sequencing; and if an appropriate translational signal is included in the primer, protein products can be made from the RNA using an in vitro translation system. Therefore, DNA and RNA probes, together with protein products and sequence information, can be obtained without time-consuming conventional cloning.

The sequences obtained in this report suggest that the AR is highly conserved among divergent species. Interestingly, a comparison of frog and human AR identified 12 silent (or synonymous) nucleotide substitutions within 32 amino acids, which implies that this domain is under very high evolutionary pressure and that changes, which would alter the amino acid sequence in this domain, are unacceptable to the survival of the species. It has been found recently that the "knuckle region" of the first zinc-finger and the amino acids immediately following this region are extremely important for the DNA binding specificity of the glucocorticoid and estrogen receptors (16). In Figure 4, this region can be observed as amino acids 557 to 570. The amino acid sequences in this "knuckle" region of the AR were identical in all the species examined. This suggests that this domain is also important for the DNA binding specificity of the AR. Therefore, by comparing the sequences obtained with the PCR technique,

functional domains of the receptors can be analyzed at both the nucleic acid and amino acid levels in different species to elucidate the evolutionary importance of a particular structure.

ACKNOWLEDGMENTS

We thank Dr. Steve S. Sommer, Dr. Gobinda Sarkar, and Dr. Cynthia D. K. Bottema for assistance in developing the RAWTS technique. This work was supported by NIH grants CA 32387, DK 37688, and NS 19949. Part of this work has been reported in abstract form at the UCLA Symposia (20).

REFERENCES

1. Chang, C., Kokontis, J., and Liao, S. (1988) Science **240**, 324-326.
2. Lubahn, D.B., Joseph, D.R., Sullivan, P.M., Willard, H.F., French, F.S., Wilson, E.M. (1988) Science **240**, 327-330.
3. Chang, C., Kokontis, J., and Liao, S. (1988) Proc.Natl.Acad.Sci.U.S.A. **85**, 7211-7215.
4. Lubahn, D.B., Joseph, D.R., Asr, M., Tan, J., Higgs, H.N., Larson, R.E., French, F.S., and Wilson, E.M. (1988) Mol.Endocrinol. **2**, 1265-1275.
5. Tan, J., Joseph, D.R., Quarmby, V.E., Lubahn, D.B., Asr, M., French, F.S., and Wilson, E.M. (1988) Mol.Endocrinol. **2**, 1276-1285.
6. Young, C.Y.F., Johnson, M.P., Prescott, J.L., Heydrick, K., Murthy, L., and Tindall, D.J. (1988) Abstract in the 1988 Annual Meeting of the Endocrinology Society, New Orleans, LA.
7. Kocher, T.D., and White, T.J. (1989) In: PCR Technology (H.A. Erlich, Ed.) pp 137-147. Stockton Press.
8. Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X., and Wilson, A.C. (1989) Proc.Natl.Acad.Sci.USA **86**, 6196-6200.
9. Sarkar, G., and Sommer, S.S. (1989) Science **244**, 331-334.
10. Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
11. Kraft, R., Tardiff, J., Krauter, K.S., and Leinwand, L.A. (1988) BioTechniques **6**, 544-546.
12. Church, G., Gilbert, W. (1984) Proc.Natl.Acad.Sci.U.S.A. **81**, 1991-1995.
13. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J., Rutter, W.J. (1979) Biochemistry **18**, 5294-5299.
14. Weiler, I.J., Lew, D., and Shapiro, D. (1987) Mol.Endocrinol. **1**, 355-362.
15. Quarmby, V.E., Yarbrough, W.G., Lubahn, D.B., French, F.S., and Wilson, E.M. (1990) Mol.Endocrinol. **4**:22-28.
16. Danielsen, M., Hinck, L., Ringold, G. (1989) Cell **57**, 1131-1138.
17. Kelley, D.B., Morrell, J.I., and Pfaff, D.W. (1975) J.Comp.Neurol. **164**, 47-62.
18. Wetzel, D., Haerter, U., and Kelley, D. (1985) J.Comp.Physiol. **A157**, 749-761.
19. Kelley, D.B., Sassoon, D., Segil, N., and Scuddel, M. (1989) Develop.Biol. **131**, 111-211.
20. He, W.W., Zhu, X.P., Young, C.F.Y., Tindall, D.T. (1990) J.Cell.Biochem., 521.