Identification of cytochrome *P*-450 1A (CYP1A) genes from two teleost fish, toadfish (*Opsanus tau*) and scup (*Stenotomus chrysops*), and phylogenetic analysis of CYP1A genes

Hilary G. MORRISON,*† Marjorie F. OLEKSIAK,* Neal W. CORNELL,‡ Mitchell L. SOGIN† and John J. STEGEMAN*§

* Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, U.S.A., † Center for Molecular Evolution, Marine Biological Laboratory, Woods Hole, MA 02543, U.S.A. and ‡ Marine Biological Laboratory, Woods Hole, MA 02543, U.S.A.

Cytochrome *P*-450-mediated responses to environmental challenges are well known in diverse animal taxa, but the evolution of the complex gene superfamily coding for these enzymes is poorly understood. Here we report a phylogenetic analysis of the cytochrome *P*-450 1A (CYP1A) genes including two new sequences determined from teleost fish, toadfish (*Opsanus tau*) and scup (*Stenotomus chrysops*). Degenerate PCR primers were used to amplify a 1.2 kbp fragment from liver cDNA. The toadfish PCR product was used as a probe to identify a full-length CYP1A clone from a toadfish liver cDNA library. The entire coding region of the scup CYP1A was obtained by rapid amplification of cDNA ends (RACE) using specific primers based on the sequence of the partial PCR product. The predicted protein sequences for toadfish and scup CYP1A shared 78% and 83% amino acid identity with rainbow trout CYP1A1 respectively. Amino acid identity with mammalian CYP1A proteins ranged from 51 to 60% for 505 aligned positions. Phylogenetic analysis of four teleost fish CYP1A genes (trout, toadfish, scup and plaice) and 12 mammalian CYP1A genes suggests a monophyletic origin of the teleost genes, with the trout gene being most divergent, and indicates three distinct groupings: mammalian 1A1, mammalian 1A2, and fish 1A. This supports the idea that the gene duplication event which gave rise to CYP1A1 and CYP1A2 occurred after the divergence of the lines leading to mammals and fish. These results establish a molecular phylogeny within the CYP1A subfamily, the first such detailed phylogenetic analysis within a cytochrome *P*-450 family.

INTRODUCTION

Members of gene families 1–4 of the cytochrome *P*-450 (CYP) superfamily [1] code for mono-oxygenases that are critically important in the activation or detoxification of lipophilic foreign compounds in animals. The function and regulation of these proteins can determine the susceptibility of species and individuals to the toxic action of drugs, environmental carcinogens, and natural products. The animal cytochrome *P*-450 genes are extraordinarily diverse, and their diversity may have arisen in part in response to the evolving repertoire of chemical defences in plants [2–4]. If so, the evolutionary divergence of the cytochrome *P*-450 gene superfamily in animals might be expected to mirror the evolutionary radiation of land plants. No phylogenetic analysis of any cytochrome *P*-450 gene subfamily has yet been published and little is known about the evolution of this superfamily despite the large number of genes that have been sequenced.

Proteins in the CYP1A subfamily are prominent in the metabolism and activation of many hydrocarbon carcinogens that are environmental contaminants [5,6] and so the occurrence and functions of CYP1A forms in diverse organisms are being investigated vigorously. There are two CYP1A forms in mammals, CYP1A1 and CYP1A2, thought to have diverged between 65 and 250 million years ago (mya) [2,7]. These proteins preferentially activate polynuclear aromatic hydrocarbon and aromatic amine procarcinogens respectively. They also metabolize endogenous compounds, including steroids [8,9]. Both CYP1A1 and CYP1A2 are induced by polynuclear and planar halogenated aromatic hydrocarbons, including the chlorinated dibenzo-*p*-dioxins, dibenzofurans and biphenyls [2,9], which are aryl hydrocarbon receptor (AhR) agonists. Mammalian CYP1A1 and CYP1A2 proteins differ in their patterns of chemical induction and in the number and location of 5′ non-coding sites identified as regulatory sequences binding the Ah receptor (xenobiotic response elements) [10–12].

Functional and immunochemical studies have shown that AhR agonists induce CYP1A proteins in many non-mammalian vertebrates [6], but sequence data to infer evolutionary relationships among the mammalian and non-mammalian CYP1As are limited. Heilmann et al. [13] cloned a CYP1A cDNA from trout liver, identifying it as CYP1A1 based on comparisons with the few mammalian sequences available at the time. Recently, two CYP1A genes have been described in trout [14], raising the possibility of a functional divergence within the CYP1A gene in teleosts comparable with the CYP1A1/CYP1A2 division in mammals. Here we describe the cloning and sequencing of additional CYP1A genes from two marine teleost fish, scup and toadfish. These species represent fish families distant from one another (Perciformes and Batrachiformes) and both distant from trout (Salmoniformes). Scup were selected because extensive information on the function and regulation of CYP1A in this

Abbreviations used: AhR, aryl hydrocarbon receptor; Arnt, AhR nuclear translocator; CIAP, calf intestinal alkaline phosphatase; CYP, cytochrome *P*-450; mya, million years ago; RACE, rapid amplification of cDNA ends; RT-PCR, reverse-transcriptase polymerase chain reaction; 1 × SSC, 150 mM NaCl/15 mM sodium citrate; STET, 100 mM Tris/HCl, pH 8.0/10 mM NaCl/1 mM EDTA/5% (v/v) Triton X-100; 1 × TAE, 40 mM Tris/acetate/1 mM EDTA; 1 × TE, 10 mM Tris/HCl, pH 7.5/1 mM EDTA.

The nucleotide sequence data reported in this paper have been submitted to GenBank and assigned accession numbers U14161 (toadfish) and U14162 (scup).

§ To whom correspondence should be addressed.

species indicates that it is extremely sensitive and responsive to environmental contaminants [15,16]. Toadfish were selected because in a study of environmental induction this species was found to be only slightly induced, even when taken from highly contaminated locations (J. W. Gooch and J. J. Stegeman, unpublished work). This species therefore presents an interesting candidate for CYP1A regulatory studies.

EXPERIMENTAL

Fish liver cDNAs

Total RNA was prepared from *Opsanus tau* (toadfish) and *Stenotomus chrysops* (scup) liver as described by Clemens [17] and poly(A)$^+$RNA was selected by chromatography on oligo(dT) cellulose. For both, cDNAs were made using 10 μg of poly(A)$^+$RNA as template for reverse transcription with Superscript II reverse transcriptase (GIBCO BRL, Gaithersburg, MD, U.S.A.) and a 15-mer oligo(dT) primer. Second-strand cDNA was synthesized with DNA polymerase I. A toadfish liver cDNA library was constructed in the plasmid vector pKK 233-2 (Pharmacia Biotech, Piscataway, NJ, U.S.A.) using *Nco*I linkers. The library contains 470000 transformants, and, in addition to the CYP1A cDNA described in this report, it has been shown to contain full-length cDNAs for 5-aminolevulinate synthase (GenBank accession number L02632) β-actin, and α-tubulin (H. L. Hellmich and N. W. Cornell, unpublished work). For reverse transcriptase-PCR (RT-PCR), poly(A)$^+$ RNA was isolated from homogenized scup liver with the FastTrack mRNA isolation kit and used with the cDNA Cycle kit, both manufactured by InVitrogen Corporation (San Diego, CA, U.S.A.).

Oligonucleotide primers

The sequences of the PCR primers initially used to amplify a 1.2 kbp fragment from both toadfish and scup cDNA templates were 5′-ctg-cag-gat-CCY-GTG-GTK-GTK-CTG-AGY-GG-3′ and 5′-aat-cga-att-CAG-CAR-GAT-GGC-CAR-GAA-GAG-RAA-3′, corresponding to amino acid positions 87–93 and 474–481 shown in Figure 3. The forward amplification primer included *Pst*I and *Bam*HI restriction sites, and the reverse primer included an *Eco*RI site. The primers included 16- and 8-fold degeneracy since many of the amino acids at the primer sites had multiple codon possibilities.

Amplification by the PCR

The primers were first used in a standard PCR reaction with 10 ng of toadfish double-stranded cDNA as a template. The reaction mixture included 30 mM Tris, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1 % (w/v) gelatin, 200 μM dNTPs, 500 ng of each primer, and 2.5 units of *Taq* polymerase. The reaction conditions were denaturation for 1 min at 92 °C, annealing for 3 min at 37 °C, and extension at 72 °C for 6 min for 30 cycles. This was followed by a single 9 min cycle at 72 °C. The products were electrophoresed through a 1 % (w/v) agarose gel in 1 × TAE buffer (40 mM Tris/acetate/1 mM EDTA) and the DNA from the predominant band at approx. 1 kbp was recovered by stabbing into the gel with a sterile micropipette tip and resuspending the agarose plug in 200 μl of 10 mM Tris (pH 7.5). The suspension was mixed vigorously by vortexing and heated to 65 °C for 10 min, then mixed again. A sample (10 μl) of this suspension was used as a template for 30 additional cycles of PCR, under the same conditions except that the annealing temperature was raised to 45 °C. The PCR product was then extracted with phenol/chloroform/isoamyl alcohol (25:24:1, by vol.), precipitated with 3 M sodium acetate, pH 5.2, and ethanol, digested with *Bam*HI and *Eco*RI, purified over a Magic miniprep column (Promega, Madison, WI, U.S.A.) and suspended in 50 μl of TE (TE: 10 mM Tris/HCl, pH 7.5/1 mM EDTA). This DNA fragment was cloned into pBluescript II KS- (Stratagene, La Jolla, CA, U.S.A.) that had been cut with *Bam*HI and *Eco*RI and treated with calf intestinal alkaline phosphatase (CIAP). The ligation products were used to transform competent cells, and DNA was prepared from transformants using a STET [100 mM Tris/HCl, pH 8.0/10 mM NaCl/1 mM EDTA/5 % (v/v) Triton X-100] boiling miniprep procedure [18].

The primers were also used to amplify an internal CYP1A fragment from scup cDNA, first using the reverse primer to synthesize single-strand cDNA from mRNA, and then using both primers in PCR as above. A specific product was seen after the first round of amplification, and the DNA was reamplified as above to yield a sufficient quantity for cloning. This DNA was cloned by blunt-end ligation into pBluescript cut with *Sma*I.

Identification and sequencing of clones

Clones containing 1.2 kbp inserts were identified by restriction enzyme digestion of miniprep DNA. Preliminary sequencing was done with both the M13 universal and reverse primers, which read into the multiple cloning sites of pBluescript in opposite orientations. Clones were then sequenced fully using CYP1A sequencing primers that were designed based on the trout CYP1A gene and on sequence information obtained from deletion clones.

Identification of full-length toadfish and scup CYP1A clones

The 1.2 kbp internal fragment of the toadfish CYP1A clone was labelled with [^{32}P-α]dCTP using the Prime-a-gene protocol (Promega, Madison, WI, U.S.A.) and used as a probe to screen the toadfish liver cDNA library. Nitrocellulose filters containing colonies were treated for 5 min each with 0.5 M NaOH/1 M Tris, pH 7.5, and 0.5 M Tris/1.25 M NaCl; then dried. The nucleic acids were immobilized on the membrane by UV cross-linking. The filters were prehybridized for 4 h at 42 °C in 50 % formamide/4 × SSC (1 × SSC: 150 mM NaCl/15 mM sodium citrate), and 0.1 % (w/v) SDS; then the radiolabelled probe was added and hybridized overnight at 42 °C. The filters were washed in 2 × SSC/0.1 % SDS three times for 10 min each and once at 65 °C for 30 min, then dried and put on to Kodak X-Omat film. One positive clone was identified, and the insert was cut out with *Nco*I and recloned into pBluescript for easier sequencing.

The 5′ and 3′ ends of the scup CYP1A gene were obtained by amplifying each from scup mRNA template in a RACE reaction [19]. The 5′ and 3′ RACE clones were sequenced, and specific primers were designed to amplify the scup CYP1A from the double-stranded scup cDNA. This product was directly cloned into the pGEM-T vector (Promega, Madison, WI, U.S.A.). Both the toadfish and scup clones were sequenced entirely.

Phylogenetic analysis of the mammalian and fish CYP1A genes

The mammalian CYP1A1 [human (*Homo sapiens*), macaque (*Macaca irus*), rabbit (*Oryctolagus cuniculus*), guinea pig (*Cavia cobaya*), rat (*Rattus norvegicus*), mouse (*Mus musculus*), and hamster (*Mesocricetus auratus*)]; mammalian CYP1A2 (human, rabbit, rat, mouse and hamster); trout (*Oncorhynchus mykiss*) and plaice (*Pleuronectes platessa*) CYP1A genes were retrieved

```
-95                                    tgcatcagactatatttgcattatttctgacagca
-60 cgttctgggcatcgcagcaccttcaggttgtgtaggatccagaaagaaggagctatcatc

    M  A  L  I  I  L  P  F  I  G  S  L  S  V  S  E  S  L  V  A
  1 ATGGCGCTAATTATACTGCCCTTCATTGGATCACTGTCAGTATCTGAGAGTCTAGTTGCC

    L  I  T  I  C  V  V  Y  L  I  L  T  Y  S  H  T  K  I  P  A
 61 TTGATAACGATATGTGTGGTGTACCTGATCCTCACGTATTCTCACACCAAGATTCCGGCA

    G  L  Q  R  L  P  G  P  K  P  L  P  I  I  G  N  V  L  E  I
121 GGCCTTCAGAGACTCCCCGGACCTAAGCCTCTACCTATCATCGGAAATGTGCTGGAGATT

    G  R  K  P  Y  Q  T  L  T  A  L  S  K  R  Y  G  P  V  F  Q
181 GGCAGGAAACCCTACCAGACCCTTACAGCATTGAGCAAGCGCTACGGACCGGTCTTCCAG

    I  Q  I  G  M  R  P  V  V  V  L  S  G  S  E  T  V  R  Q  A
241 ATTCAGATCGGCATGCGTCCTGTGGTCGTGCTGAGTGGCAGTGAGACAGTTCGGCAGGCT

    L  I  K  Q  G  E  D  F  S  G  R  P  D  L  Y  T  F  Q  F  I
301 CTTATCAAGCAAGGGAAGATTTTTCAGGTAGACCTGACCTGTACACCTTTCAGTTCATC

    S  D  G  K  S  L  A  F  S  T  D  K  V  G  V  W  R  A  R  R
361 AGTGACGGCAAGAGTCTGGCCTTCAGCACAGACAAGGTCGGTGTATGGCGCGCCCGCCGC

    K  L  A  Y  S  A  L  R  S  F  S  S  L  E  S  T  N  Q  E  Y
421 AAGCTAGCCTACAGCGCTCTGCGCTCCTTCTCCAGCCTGGAGAGCACGAACCAGGAGTAC

    S  C  M  L  E  E  H  I  C  K  E  G  E  Y  L  V  K  Q  L  N
481 TCCTGCATGCTGGAGGAGCACATCTGCAAAGAGGGAGAGTATCTGGTGAAACAGCTGAAC

    T  S  M  K  A  N  G  S  F  D  P  F  R  N  I  V  V  S  V  A
541 ACCTCGATGAAGGCCAACGGCAGCTTCGACCCGTTCCGCAACATTGTGGTCTCTGTTGCA

    N  V  I  C  G  M  C  F  G  R  R  Y  D  H  Y  D  Q  E  L  V
601 AACGTGATCTGCGGGATGTGCTTCGGCCGACGCTACGACCATTACGACCAGGAGCTGGTC

    S  L  V  N  L  S  E  E  F  G  Q  V  V  G  T  G  N  L  A  D
661 AGCTTGGTGAACCTCAGTGAGGAGTTCGGCCAGGTGGTGGGAACAGGAAACCTAGCAGAC

    F  I  P  V  L  R  F  L  P  S  T  A  M  K  K  F  L  S  I  N
721 TTCATCCCTGTTCTCCGTTTCCTGCCCAGCACGGCGATGAAGAAATTTTTGAGCATCAAT

    D  R  F  D  K  F  V  K  K  I  V  S  E  H  Y  A  T  Y  N  K
781 GATCGTTTTGACAAGTTTGTGAAGAAGATAGTCAGCGAGCACTATGCCACATATAACAAG

    D  N  I  R  D  I  T  D  S  L  I  D  H  C  E  D  R  K  L  D
841 GACAACATTCGTGACATCACCGACTCCCTCATCGATCACTGTGAGGACAGGAAGCTGGAC

    E  N  C  N  V  Q  V  S  D  E  K  I  V  G  I  V  N  D  L  F
901 GAAAACTGCAACGTCCAGGTGTCAGATGAAAAGATTGTAGGAATCGTCAATGACCTGTTT

    G  A  G  F  D  T  V  S  T  G  L  S  W  S  V  M  Y  L  V  A
961 GGCGCTGGTTTTGACACCGTCTCCACCGGTCTGTCATGGTCTGTGATGTACTTAGTGGCT

    Y  P  E  I  Q  E  R  L  Y  Q  E  I  K  D  S  V  G  T  E  R
1021 TATCCAGAGATACAGGAAAGGCTTTATCAAGAAATTAAAGACAGTGTGGGAACAGAGCGC

    M  P  L  L  S  D  R  P  S  L  P  F  L  D  A  F  I  L  E  I
1081 ATGCCTCTTCTCTCTGACAGACCCAGCTTGCCTTTTCTGGATGCCTTTATCCTGGAGATC

    F  R  H  S  S  F  L  P  F  T  I  P  H  C  T  T  K  N  T  S
1141 TTTAGACACTCTTCATTCCTGCCATTCACCATCCCACACTGCACGACAAAAAACACATCT

    L  N  G  Y  F  I  P  K  D  T  C  V  F  I  N  Q  W  Q  I  N
1201 CTTAATGGCTACTTCATCCCTAAAGACACCTGCGTGTTCATCAATCAGTGGCAGATCAAC

    H  D  P  E  L  W  K  D  P  F  S  F  N  P  E  R  F  L  S  A
1261 CATGATCCTGAGCTGTGGAAAGATCCGTTCTCCTTCAACCCTGAGCGCTTCCTGAGCGCC

    D  G  T  E  L  N  R  L  E  G  E  K  V  M  L  F  G  L  G  K
1321 GACGGCACTGAGCTCAACAGGCTAGAAGGGGAGAAGGTGATGCTTTTCGGCTTGGGCAAA

    R  R  C  I  G  E  V  I  A  R  N  E  V  F  L  F  L  A  I  I
1381 CGGCGATGCATCGGTGAGGTCATCGCCAGAAACGAAGTCTTCCTCTTCTTGGCAATCATC

    I  Q  R  L  Q  F  H  M  L  P  G  E  P  L  D  M  T  P  E  Y
1441 ATCCAGAGGCTTCAGTTCCACATGCTGCCTGGAGAGCCCTTGGACATGACGCCGGAATAC

    G  L  T  M  K  H  K  R  C  Q  L  R  A  T  M  R  E  K  N  E
1501 GGCCTCACCATGAAGCACAAACGCTGCCAGCTGAGGGCTACCATGAGAGAGAAGAATGAG

     Q  *
1561 CAGTGAggcccgttaataaatacgctgcgtgggtcgcagtcacatgaagaggcaacgaat
1621 tctcacacttaaggcaacgggaagcaaactggatttctgcggctgcttagactgaatttg
1681 tctgggtttttttcttcactgagagacactcatgcctctcgttgaagaatcaggagagcc
1741 tcttttctgctaagctagactgccattcagaattctcaaactgcagcaggcgattgtgcg
1801 aggttaagttgccttgttctgctttgttttgctggataggcgttgctgtttaggacactg
1861 cagccaacagttttgtgtgatgttgaactctcaactcaaccgaaatggtgattttacttt
1921 atgaacacctttacacatggaagctaattattgtaattccattctactgtggatgttaa
1981 acactgaaacggtcccaaaatgtgatataatacatatatgtatcagatttatgcattcga
2041 tgtgccttgtatgtgtacataattgaggattaagtgtatttgctgtgcatttatttgatt
2101 tgtaaggaaatcaaaaccattttgttgatgtttttgaatttatggtgacatgatattgta
2161 ttttttatcacttatccatccatgaatgtatcctatataaaacaaacttgtgatcaaag
2221 tatattttaatgactcaccaataaaatataaaatacaaaaaaaaaa
```

Figure 1 Toadfish CYP1A

The nucleotide sequence is 2361 bp long and encodes a predicted protein of 521 amino acids. The gene was obtained from a toadfish liver cDNA library.

```
    M  V  L  M  I  L  P  V  I  G  S  V  S  V  S  E  G  L  V  A
  1 ATGGTGCTGATGATATTGCCAGTCATTGGATCGGTGTCGGTGTCGGAGGGTTTGGTGGCC

    M  I  T  M  C  L  A  Y  L  I  L  R  L  F  R  T  E  I  P  E
 61 ATGATAACCATGTGTCTGGCCTACCTGATCCTCAGGCTTTTCCGCACTGAGATTCCCGAG

    G  L  L  Q  L  P  G  P  K  P  L  P  I  I  G  N  V  L  E  V
121 GGGCTTCTTCAGCTGCCCGGACCGAAGCCCCTGCCCATCATTGGGAATGTGCTCGAGGTG

    G  R  N  P  Y  L  S  L  T  A  M  S  K  R  Y  G  D  V  F  Q
181 GGAAGAAATCCTTACCTGAGTCTCACTGCCATGAGCAAGCGCTACGGCGACGTCTTCCAG

    I  Q  I  G  M  R  P  V  V  V  L  S  G  S  E  T  V  R  Q  A
241 ATCCAGATTGGCATGCGTCCCGTCGTCGTGTTGAGCGGCAGTGAAACAGTCCGACAGGCT

    L  I  K  Q  G  D  D/E  F  A  G  R  P  D  L  Y  S  F  R  F  I
301 CTCATCAAGCAAGGGGACGASTTTGCAGGCAGGCCTGACCTGTACAGCTTCAGGTTTATC

    N  D  G  K  S  L  A  F  S  T  D  Q  A  G  V  W  R  A  R  R
361 AACGATGGCAAGAGTCTGGCCTTCAGCACGGACCAGGCCGGCGTGTGGCGTGCCCGCAGG

    K  L  A  Y  S  A  L  R  S  F  A  T  L  E  G  T  T  P  E  Y
421 AAGCTGGCCTACAGTGCCCTGCGCTCCTTYGCCACCTTGGAAGGCACGACGCCAGAGTAC

    S  C  A  L  E  E  H  V  S  K  E  A  E  Y  L  V  K  Q  L  H
481 TCCTGTGCGCTGGAGGAACACGTCAGCAAAGAGGCAGAGTATCTGGTCAAACAGTTGCAC

    T  V  M  E  A  D  G  S  F  D  P  F  R  H  I  V  V  S  V  A
541 ACCGTCATGGAGGCCGATGGCAGCTTCGACCCCTTCCGCCACATTGTCGTCTCCGTCGCT

    N  V  I  C  G  M  C  F  G  R  R  Y  D  H  N  H  Q  E  L  L
601 AACGTGATCTGTGGCATGTGCTTCGGCCGACGCTACGACCACAACCACCAGGAGCTGCTC

    N  L  V  N  L  S  D  E  F  G  Q  V  V  A  S  G  N  P  A  D
661 AACTTGGTGAACCTCAGCGACGAGTTCGGCCAGGTGGTGGCCAGCGGTAACCCTGCRGAC

    F  I  P  I  L  Q  Y  L  P  S  T  T  M  K  K  F  L  N  I  N
721 TTCATYCCTATCCTYCAGTACCTGCCCAGCACAACRATGAAGAAGTTTTTGAACATCAAC

    D  R  F  N  T  F  V  Q  K  I  V  S  E  H  Y  T  T  F  D  K
781 GACCGCTTCAACACGTTTGTGCAAAAGATCGTCAGCGAGCACTACACCACCTTTGACAAG

    D  N  I  R  D  I  T  D  S  L  I  D  H  C  E  D  R  K  L  D
841 GACAACATTCGGGACATCACAGACTCCCTCATTGATCACTGCGAGGACAGGAAGCTGGAT

    E  N  S  N  V  Q  M  S  D  E  K  I  V  G  I  V  N  D  L  F
901 GAGAACTCCAATGTCCAGATGTCGGATGAGAAGATTGTAGGAATTGTCAATGACCTGTTT

    G  A  G  F  D  T  I  S  T  A  L  S  W  S  V  M  Y  L  V  A
961 GGAGCTGGTTTTGACACCATCTCTACTGCCCTGTCTTGGTCTGTGATGTACTTGGTTGCA

    Y  P  E  I  Q  E  R  L  Y  Q  E  M  N  E  T  V  G  P  D  R
1021 TACCCAGAGATACAGGAAAGGCTTTATCAAGAAATGAACGAAACTGTGGGCCCRGATCGC

    T  P  C  L  S  D  K  P  K  L  P  F  L  E  A  F  I  L  E  T
1081 ACGCCTTGTCTCTCTGACAAACCCAAGTTACCCTTCCTGGAGGCCTTCATCCTGGAAACC

    F  R  H  S  S  F  L  P  F  T  I  P  H  C  T  S  K  D  T  S
1141 TTTCGCCATTCTTCATTCCTGCCCTTCACCATCCCTCACTGCACATCAAAAGACACATCT

    L  N  G  Y  F  I  P  K  D  T  C  V  F  I  N  Q  W  Q  I  N
1201 CTGAACGGCTACTTCATTCCCAAAGACACCTGTGTCTTCATCAATCAGTGGCAGATCAAC

    H  D  A  E  L  W  K  D  P  S  S  F  N  P  D  R  F  L  N  A
1261 CATGATGCTGAGCTGTGGAAAGATCCGTCTTCCTTCAACCCAGACCGCTTCCTGAATGCG

    D  G  T  E  V  N  K  L  E  G  E  K  M  M  V  F  G  M  G  K
1321 GACGGCACCGAGGTCAACAAGCTAGAGGGGGAGAAGATGATGGTGTTCGGCATGGGAAAG

    R  R  C  I  G  E  V  I  A  R  S  E  V  F  L  F  L  A  I  L
1381 CGACGCTGCATCGGAGAGGTCATTGCAAGAAGTGAAGTCTTCCTCTTCTTGGCGATTCTC

    V  Q  N  L  R  F  H  S  M  P  G  E  P  L  D  M  T  P  E  Y
1441 GTCCAGAATCTGCGGTTCCACTCGATGCCCGGAGAGCCACTGGACATGACCCCGGAATAC

    G  L  T  M  K  H  K  R  C  Q  L  R  A  A  M  R  A  R  N  E
1501 GGTCTCACAATGAAGCACAAACGCTGCCAACTGAGAGCCGCGATGCGAGCACGCAATGAG

     E  *
1561 GAGTGA
```

Figure 2 Scup CYP1A

The coding region is 1566 bp long and encodes a predicted protein of 521 amino acids. The gene was obtained by PCR amplification from scup liver cDNA.

from the GenBank database (see references in Figure 4). The scup and toadfish genes were aligned with these published sequences using CLUSTAL [20]; refinements to the alignment were made manually. Phylogenetic trees were constructed using distance-matrix [21] and maximum parsimony methods [22].

RESULTS

Nucleotide and inferred amino acid sequences of scup and trout CYP1A

Opsanus tau and *Stenotomus chrysops* CYP1A partial cDNAs were amplified using double-stranded cDNA and RT-PCR templates respectively. The primary PCR product obtained with each species template was 1185 bp. These products were cloned into the plasmid vector pBluescript II KS- for sequencing. A single clone containing toadfish cDNA fragment and three clones containing scup cDNA were sequenced.

The toadfish clone was used as a probe to identify a full-length cDNA clone from a toadfish liver cDNA library by colony hybridization. The toadfish CYP1A gene was 2.4 kbp, including 5′ and 3′ non-coding domains. The coding region was 1566 bp. The complete cDNA sequence and predicted protein sequence are shown in Figure 1.

Internal scup CYP1A-specific primers were used to amplify the 5′ and 3′ ends of the gene from scup liver mRNA in separate RACE reactions. The PCR products were cloned and sequenced. Primers corresponding to conserved sites at the 5′ and 3′ termini were used to amplify the full-length coding sequence from scup liver cDNA, which was cloned into the pGEM vector by T-A ligation. A single clone and a pool of five additional clones were sequenced. The coding region of the scup CYP1A gene was 1566 bp long, and the consensus sequence is shown in Figure 2. Sequence ambiguity occurred at only seven positions and six of the nucleotide changes were at third positions corresponding to silent mutations. At nucleotide 321, the third position variation produced an amino acid ambiguity, glutamic acid versus aspartic acid. The origin of these ambiguities is uncertain, and it is possible that errors in the gene could have arisen during the PCR amplification. However, the Glu-Asp ambiguity was observed in the products of two separate PCR reactions and ambiguities occurred only at third positions, suggesting that they occur in the mRNA population rather than being generated by *Taq* polymerase error, and thus could represent alleles. The amino acid sequence of the 5′ end of the scup CYP1A1 protein has been determined by protein sequencing [23] and is identical to the predicted protein sequence of the scup CYP1A gene.

Table 1 Nucleotide (a) and amino acid (b) conservation between aligned CYP1A genes

(a) Proportion nucleotide identities were calculated for the 1515 aligned positions using the Sequence Analysis Software Package of the Genetics Computing Group, Inc. (GCG, Madison, WI, U.S.A.). (b) Proportion amino acid identities were calculated for the 505 aligned positions using the Sequence Analysis Software Package (GCG).

(a)	Monkey1A1	Human1A1	Rabbit1A1	Mouse1A1	Rat1A1	Hamster1A1	Guinea pig1A1	Human1A2	Rabbit1A2	Mouse1A2	Rat1A2	Hamster1A2	Trout1A	Scup1A	Toadfish1A	Plaice1A
Monkey1A1	1.0000	0.9524	0.8155	0.8254	0.8181	0.8234	0.7989	0.7913	0.7508	0.7405	0.7412	0.7392	0.6323	0.6356	0.6177	0.6283
Human1A1		1.0000	0.8194	0.8261	0.8234	0.8267	0.8056	0.7952	0.7522	0.7365	0.7432	0.7405	0.6376	0.6402	0.6177	0.6316
Rabbit1A1			1.0000	0.8009	0.7970	0.7970	0.7817	0.7740	0.8184	0.7265	0.7232	0.7132	0.6296	0.6369	0.6124	0.6257
Mouse1A1				1.0000	0.9286	0.8929	0.7937	0.7508	0.7223	0.7651	0.7538	0.7359	0.6250	0.6369	0.6151	0.6263
Rat1A1					1.0000	0.8882	0.7917	0.7462	0.7177	0.7425	0.7598	0.7319	0.6237	0.6323	0.6104	0.6217
Hamster1A1						1.0000	0.7844	0.7462	0.7131	0.7325	0.7352	0.7545	0.6118	0.6224	0.5979	0.6151
Gpig1A1							1.0000	0.7488	0.7290	0.7106	0.7226	0.7092	0.6270	0.6270	0.5992	0.6276
Human1A2								1.0000	0.8363	0.8017	0.8057	0.8044	0.6289	0.6382	0.6090	0.6163
Rabbit1A2									1.0000	0.7838	0.7838	0.7645	0.6243	0.6276	0.6030	0.5971
Mouse1A2										1.0000	0.9341	0.8856	0.6075	0.6234	0.6001	0.6048
Rat1A2											1.0000	0.8882	0.6075	0.6261	0.5995	0.6061
Hamster1A2												1.0000	0.6021	0.6121	0.5848	0.5975
Trout1A													1.0000	0.7908	0.7822	0.7729
Scup1A														1.0000	0.8244	0.8581
Toadfish1A															1.0000	0.8158
Plaice1A																1.0000

(b)

	Monkey1A1	Human1A1	Rabbit1A1	Mouse1A1	Rat1A1	Hamster1A1	Guinea pig1A1	Human1A2	Rabbit1A2	Mouse1A2	Rat1A2	Hamster1A2	Trout1A	Scup1A	Toadfish1A	Plaice1A
Monkey1A1	1.0000	0.9365	0.7698	0.8036	0.7937	0.7857	0.7817	0.7296	0.6918	0.6826	0.6766	0.6647	0.5675	0.5734	0.5496	0.5655
Human1A1		1.0000	0.7679	0.8036	0.7956	0.7956	0.7857	0.7276	0.6859	0.6806	0.6747	0.6627	0.5655	0.5774	0.5516	0.5694
Rabbit1A1			1.0000	0.7718	0.7718	0.7679	0.7401	0.7097	0.7296	0.6467	0.6427	0.6287	0.5675	0.5694	0.5575	0.5675
Mouse1A1				1.0000	0.9325	0.8810	0.7817	0.6899	0.6759	0.7226	0.7126	0.6926	0.5893	0.5992	0.5754	0.5833
Rat1A1					1.0000	0.8770	0.7798	0.6839	0.6700	0.6886	0.7026	0.6707	0.5913	0.5972	0.5714	0.5873
Hamster1A1						1.0000	0.7599	0.6799	0.6640	0.6707	0.6766	0.6866	0.5615	0.5754	0.5556	0.5714
Gpig1A1							1.0000	0.6918	0.6740	0.6587	0.6747	0.6507	0.5893	0.5952	0.5675	0.5873
Human1A2								1.0000	0.7873	0.7325	0.7545	0.7345	0.5328	0.5527	0.5229	0.5388
Rabbit1A2									1.0000	0.7445	0.7405	0.7166	0.5268	0.5348	0.5189	0.5189
Mouse1A2										1.0000	0.9301	0.8842	0.5190	0.5289	0.5130	0.5150
Rat1A2											1.0000	0.8802	0.5130	0.5269	0.5090	0.5170
Hamster1A2												1.0000	0.5090	0.5230	0.5050	0.5150
Trout1A													1.0000	0.8337	0.7802	0.7980
Scup1A														1.0000	0.8099	0.8554
Toadfish1A															1.0000	0.8218
Plaice1A																1.0000

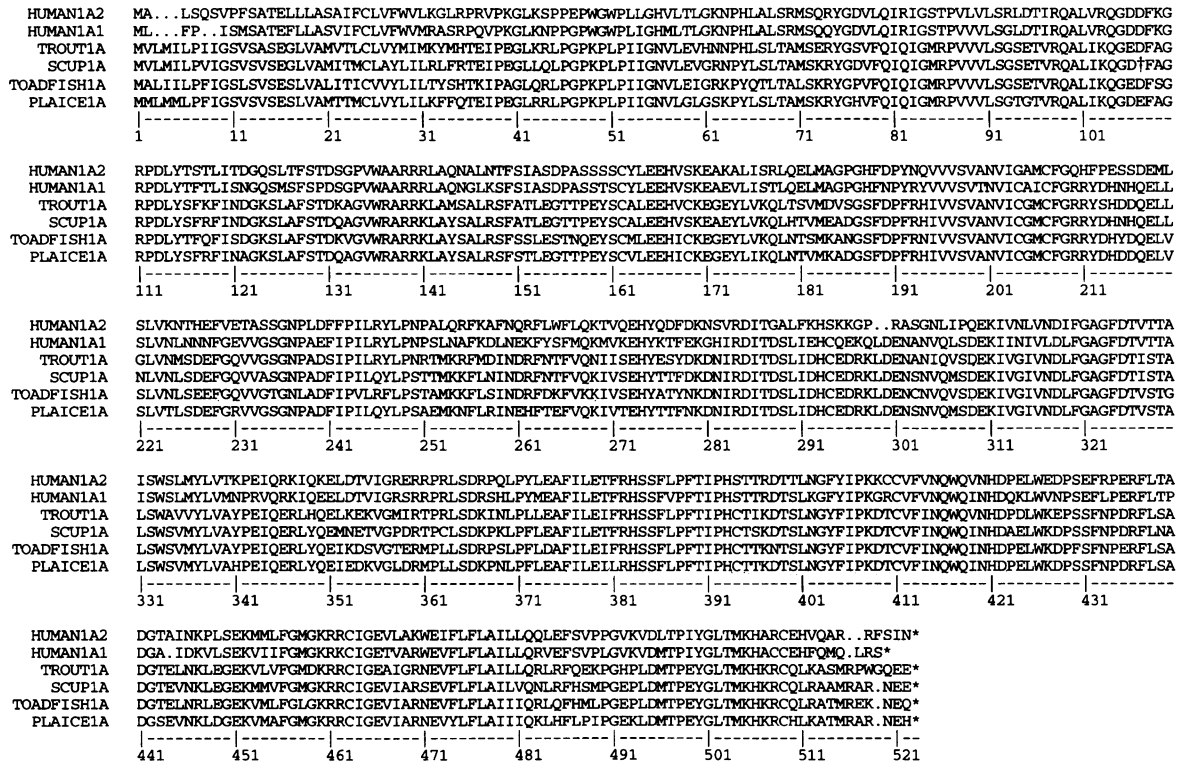

Figure 3 Alignment of human CYP1A1, human CYP1A2, and fish CYP1A protein sequences

Numbering is based on the trout CYP1A sequence. Site indicated by † in the scup protein is variable, Asp (D) or Glu (E).

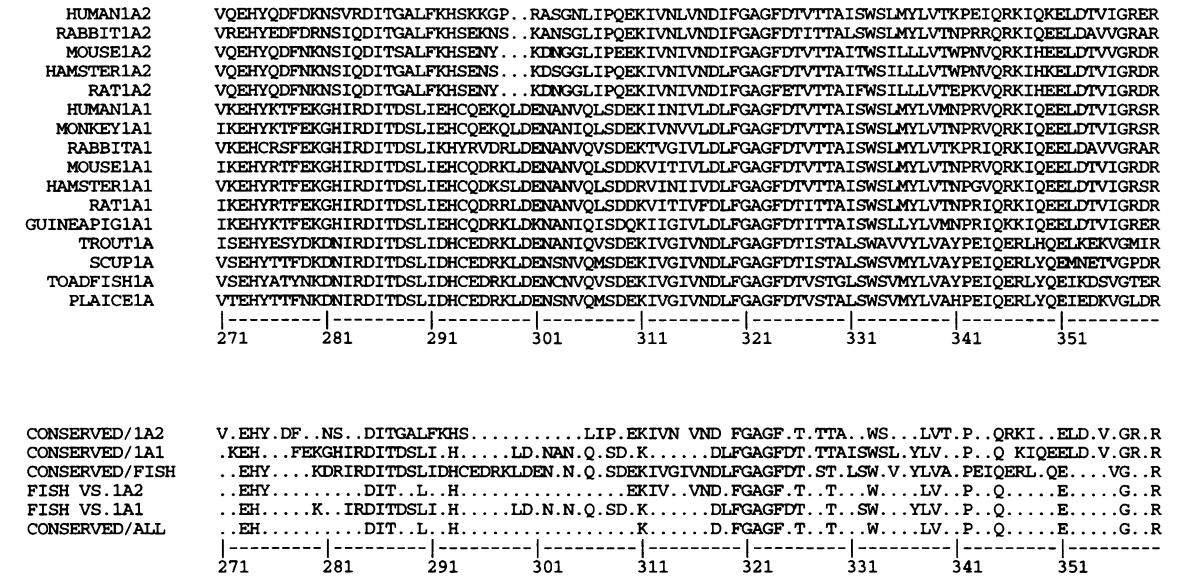

Figure 4 Alignment of mammalian CYP1A1, mammalian CYP1A2, and fish CYP1A protein sequences between amino acids 271 and 360

Numbering is based on the trout CYP1A sequence. The conservation of amino acid residues between the different groups are shown at the bottom of the Figure. Sequences were retrieved by GenBank accession numbers M55053 (human CYP1A2), D00213 (rabbit CYP1A2), X00479 (mouse CYP1A2), M34446 (hamster CYP1A2), K02422 (rat CYP1A2), K03191 (human CYP1A1), D17575 (monkey CYP1A1), D00212 (rabbit CYP1A1), Y00071 (mouse CYP1A1), D12977 (hamster CYP1A1), K02246 (rat CYP1A1), D11043 (guinea pig CYP1A1), M21310 (trout CYP1A) and X73631 (plaice CYP1A).

The nucleotide sequences were aligned with other vertebrate CYP1A1 and CYP1A2 genes available from GenBank. A total of 1515 positions out of 1569 were compared and at least one nucleotide change occurred at 855 of these positions. Length variation occurred at both the 5′ and 3′ ends relative to the mammalian genes. These regions could not be aligned and were

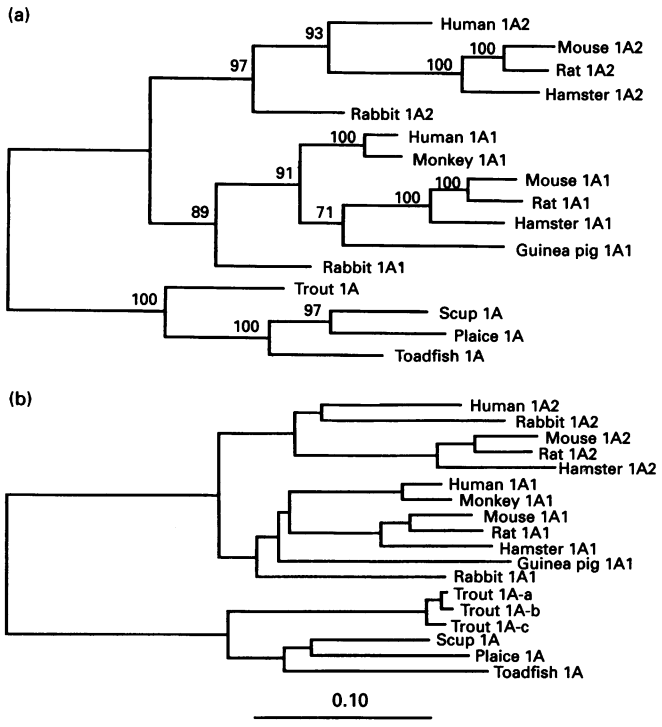

Figure 5 Phylogeny of the CYP1A genes

The tree inferred by the maximum parsimony method using a heuristic procedure is shown in (**a**). Numbers above branches are bootstrap values based on 100 samplings and are a measure of relative confidence in an assembly. The tree inferred by the distance method is shown in (**b**). Distances between genes are represented by the sum of their horizontal separation. The scale bar under the tree indicates the distance corresponding to 10 differences in 100 positions. Trout CYP1A-a is from Heilmann et al. [13]; trout CYP1A-b is 'CYP1A2' from Berndtson and Chen [14]; trout CYP1A-c is 'CYP1A1' from Berndtson and Chen [14].

not included in the data set. The conservation of nucleotide sequence among the CYP1A genes from teleost fish is high (Table 1a), ranging from 77% identity between trout and plaice to 86% between scup and plaice. There was 78% and 79% identity of the trout CYP1A1 with the toadfish and scup CYP1A1 respectively. In contrast, the similarity between the mammalian CYP1A1/CYP1A2 genes and the toadfish and scup CYP1A genes ranged from 58 to 64%.

Figure 3 presents the toadfish and scup predicted protein sequences in an alignment that includes the trout CYP1A, plaice CYP1A, human CYP1A1, and human CYP1A2 for reference. The toadfish and scup predicted protein sequences shared 78% and 83% amino acid identity with the trout protein (Table 1b). The toadfish and scup predicted proteins share 50–58% amino acid identity with the mammalian CYP1A1 and CYP1A2s. The overall amino acid identity values are slightly higher between the fish CYP1As and the mammalian CYP1A1s (average 57%) than between fish CYP1As and mammalian CYP1A2s (average 52%).

Close inspection of the alignment (Figure 4) reveals some domains that are conserved among all the CYP1A proteins. The region of highest conservation is seen for amino acids 320 to 326 (Phe-Gly-Ala-Gly-Phe-Asp-Thr). In contrast, at amino acids 281–282 all three groups can be distinguished, by Asn-Ser in mammalian CYP1A2, Gly-His in mammalian CYP1A1, and Asp-Asn in fish CYP1A.

Other regions show fish sequences that resemble mammalian CYP1A1 much more that CYP1A2. For example, between amino acids 271 to 360, there are 17 amino acids that are conserved only between mammalian CYP1A1s and fish CYP1As, but only six are conserved exclusively between mammalian CYP1A2s and fish CYP1As. At the remaining 67 sites, 24 residues are conserved between mammalian CYP1A1s, mammalian CYP1A2s, and fish CYP1As and 43 are variant between the three groups. From amino acids 283 to 291 and from 299 to 309 there is clear similarity between mammalian CYP1A1s and fish CYP1A. In contrast, only short regions in which the fish CYP1As are similar to the mammalian CYP1A2s can be identified, such as at amino acid 317, which is Asn in mammalian CYP1A2s and fish CYP1As, but variable in mammalian CYP1A1s; and at amino acids 310–313, which are Glu-Lys-Ile-Val in CYP1A2 and fish CYP1As but variable in CYP1A1.

Phylogenetic analyses of CYP1A1 and CYP1A2 genes

Over 96% (1515 nucleotides) of the sequence alignment was used to infer the phylogenetic relationships between the mammalian and fish CYP1A genes. Trees were generated using both maximum parsimony and distance methods, and similar topologies were obtained for both methods (Figures 5a and 5b).

Three distinct monophyletic branches were observed: mammalian CYP1A1, mammalian CYP1A2 and fish CYP1A. The branch order of the mammalian CYP1As differs slightly between the two trees; the distance algorithm indicates a monophyletic origin for the human and rabbit CYP1A2 genes within this group, whereas the parsimony algorithm generates separate branches. The distance tree also shows the guinea pig CYP1A1 branching outside of the primate genes, thus making the rodent genes polyphyletic. Graur et al. [24] have proposed that the guinea pigs and myomorphs (rat-like rodents) do not form a monophyletic group, which is consistent with our results. The parsimony tree shows the rodents as monophyletic, but only 71% of the bootstrap replicates support this assembly. The variable branch order within the mammalian subgroups is consistent with other studies which have failed to establish consistent branching patterns for eutherian mammalian groups, including primates, carnivores, rodents and lagomorphs [25–28].

The branch order within the fish CYP1As agrees with what is known of the evolution of teleosts; with trout diverging prior to the separation of scup and plaice. Two distinct trout CYP1A genes recently described were named CYP1A1 and CYP1A2 [14]. When these trout gene sequences were included in the alignment and phylogenetic analysis (Figure 5b), both of these genes fall within the fish CYP1A branch, with the 'CYP1A2' appearing more like the trout CYP1A1 gene reported by Heilmann et al. [13] and the 'CYP1A1' appearing to be slightly more distantly related. The inferred trees do not support an orthologous relationship between the 'CYP1A2' gene and the mammalian CYP1A2 genes. All the reported trout CYP1A genes appear to be very closely related homologues.

The amino acid conservation between trout and scup is 83% (Table 1b), which implies a closer relationship between these two species than is apparent in the phylogenetic tree based on the nucleotide sequence. Parsimony trees constructed using either the nucleotide alignment or the protein alignment vary in the placement of scup CYP1A. Figure 6 shows the teleost CYP1A phylogenetic trees inferred using all codon positions (Figure 6a), first and second positions (Figure 6b) and amino acid sequence (Figure 6c). Only the tree generated from all positions is strongly supported, with bootstrap values of 100 for all groupings. The position of the scup CYP1A gene cannot be determined in the

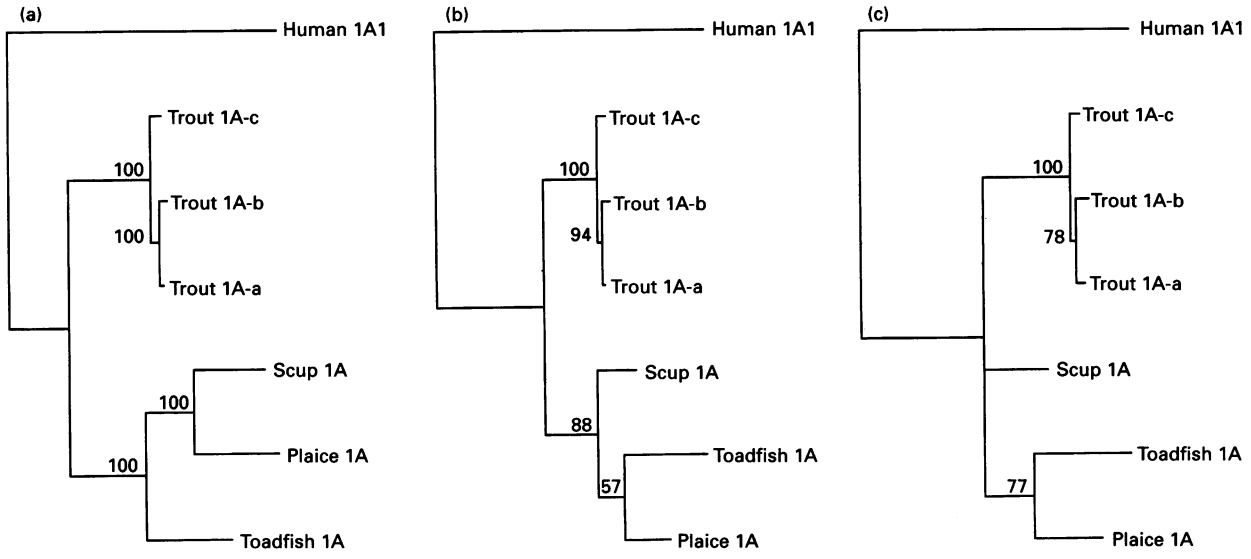

Figure 6 Phylogeny of the fish CYP1A genes

Phylogenetic trees inferred by the maximum parsimony method using the branch-and-bound procedure for the fish CYP1A genes using (**a**) all codon positions, totalling 1515 bases, (**b**) first and second codon positions included, totalling 1010 bases, (**c**) amino acids, totalling 505 residues. Numbers above branches indicate bootstrap values. Identification of trout genes as in Figure 5.

tree constructed from the amino acid sequence. The scup gene apparently is evolving more slowly (has a slower 'molecular clock') than the toadfish and plaice genes; this is indicated on the distance tree (Figure 5b) by the shorter branch length of scup compared with toadfish and plaice.

DISCUSSION

The CYP1A genes of scup and toadfish appear to be very closely related to previously reported trout and plaice CYP1A1 genes. The phylogenetic analysis shows that the fish CYP1A genes are orthologous to both the mammalian CYP1A1 and CYP1A2 subfamilies. Both trees support the hypothesis that the gene duplication event that led to the presence of both CYP1A1 and CYP1A2 in mammals occurred after the phylogenetic split between fish and the line leading to mammals. It has been suggested that the CYP1A1 and CYP1A2 genes diverged from one another not more than 250 mya [2,7], while the fish and mammalian lines are thought to have diverged more than 400 mya, prior to that split in the CYP1A gene line. However, without additional vertebrate CYP1A sequences (e.g. from birds or reptiles) it is not possible to define the time of the CYP1A divergence more precisely. Multiple proteins that differ in their N-terminal amino acid sequence are induced by AhR agonists in birds [29,30], but the relationship of these avian forms to the mammalian CYP1A proteins has not been established.

Evidence from both Northern- and Southern-blot analyses of trout [13], scup and *Fundulus heteroclitus* [31], and plaice [32] suggests that only one form of the CYP1A gene is present in fish. The concept of a single ancestral form that evolved into the modern CYP1A gene in teleosts [13,33] is supported by the presence in the fish proteins of some amino acids specific to CYP1A1 and others specific to CYP1A2. The trout gene initially sequenced was classified as CYP1A1 based on degree of sequence identity with mammalian CYP1A1 versus CYP1A2. By that criterion, all of the teleost CYP1As would be classified as CYP1A1. However, the proportion of shared nucleotide or amino acid residues should not be the only basis for such an assignment.

It has been demonstrated that the scup and other teleost CYP1A proteins are CYP1A1-like in several properties, including low-spin haem iron, catalytic functions and patterns of induction, particularly the strong induction in extrahepatic organs (see reference [6] for review). Antigenic similarities also suggest a closer relationship between teleost CYP1A and mammalian CYP1A1 proteins [33]. Examination of the amino acid alignment reveals regions of 10 or more amino acids in fish CYP1A proteins that resemble the corresponding domains in the mammalian CYP1A1s, whereas regions that resemble CYP1A2 are only 1 or 2 amino acids long. The region from 283 to 291 is a stretch of amino acids invariant between CYP1A1s and all fish CYP1As, though with a core of three amino acids found also in CYP1A2. This site appears to be antigenic, and antibodies to peptides containing that sequence distinguish between CYP1A1 and CYP1A2 proteins [34]. A monoclonal antibody to scup CYP1A strongly recognizes all of the CYP1A1s and fish CYP1As examined, but binds weakly or not at all to CYP1A2s [33,35]. The hypothesis that this monoclonal antibody binds to the 283–291 region is currently under investigation; this region could define a CYP1A1 signature valid across the vertebrata. Thus, both sequence information and studies of function to date suggest that the fish CYP1A genes represent the CYP1A1 line, and that the mammalian CYP1A2 diverged from this line. Despite the evidence suggesting that fish CYP1As are more like CYP1A1 than like CYP1A2 in function, it may be preferable to use the name CYP1A. This designation would indicate their evolution from the single ancestral CYP1A precursor gene.

Although evidence to date points to but a single CYP1A gene in most fishes, recently there have been two CYP1A genes reported in trout, designated CYP1A1 and CYP1A2 [14]. These genes are very closely related, being greater than 96% identical in the coding region. The two proteins differ most at the C-terminus, resulting from a length difference of the inferred proteins (522 versus 536 amino acids). This length difference

could have arisen from a single nucleotide deletion in codon 516 (Arg) of clone 10 (the CYP1A2); an alignment gap introduced at this position results in identical 3′ amino acid sequences and termination sites for the two genes. The appearance and the source of a deletion that might contribute to such a dramatic change in the 3′ end of the protein is intriguing.

The two trout genes were designated CYP1A1 and CYP1A2, apparently based on the presence of xenobiotic response elements in the proximal 5′ non-coding region of clone 1 (the CYP1A1) and the absence of these elements in the same region of clone 10 (the CYP1A2). The authors suggest that the trout CYP1A2 is not orthologous to the mammalian CYP1A2 gene and suggest that a separate duplication event occurred. The analysis shown in Figure 5(a) establishes that the trout CYP1A1 and CYP1A2 cannot be orthologous to the mammalian CYP1A1 and CYP1A2 respectively, raising questions about the terminology. Interestingly, our phylogenetic analysis shows that the sequence of the trout CYP1A1 reported by Heilmann et al. [13] is more similar to the trout 'CYP1A2' (clone 10) than 'CYP1A1' (clone 1). The relationship between these genes and their relationship to other CYP1A genes needs clarification and their designation as CYP1A1 and CYP1A2 bears reconsideration. An alternative designation of teleost CYP1 genes in a CYP1B subfamily would not be appropriate, since a human CYP1B1 gene was recently reported ([36]; GenBank accession number U03688), which is quite different from the CYP1As. Designation of such a new CYP1A gene as CYP1A3 consistent with earlier suggestions [33], may be more appropriate; however, all three genes contain the identical CYP1A1 sequence at amino acid positions 283–291. We would agree with a terminology of CYP1A3 for clone 1 of Berndtson and CYP1A1 for clone 10.

The occurrence of two genes in trout appears to be quite recent. Just as CYP2B1 and CYP2B2 are recently diverged CYP2B forms found only in rat, the two forms of CYP1A identified in trout are likely to be specific to salmonids or even a single genus. We suggest that counterparts to only one of these genes will be found in other fishes. That does not preclude, however, the occurrence of multiple CYP1A-like proteins derived from divergence of genes in other fish families or species. Clearly, additional CYP1A sequences, including regulatory regions, are required from other species to properly address these questions of multiplicity.

The CYP1A genes and their regulatory partners, the AhR and the AhR nuclear translocator (Arnt) proteins, appear to have undergone fundamental changes soon after vertebrate emergence [37]. Whether this resulted from the acquisition of new biological functions is not yet clear. Additional sequences in lower vertebrates will be important to addressing that question, and in establishing which groups are susceptible to dioxin and dioxin-like compounds acting through the AhR and to other compounds that are substrates for CYP1A proteins.

The research was supported in part by NIH ES05588 to H.G.M., ES04220 to J.J.S., and GM32964 to M.L.S. We thank Grace Bruning of the Marine Biological Laboratory for technical assistance. Contribution number 8782 from the Woods Hole Oceanographic Institution.

Received 6 September 1994/5 December 1994; accepted 9 December 1994

REFERENCES

1 Nelson, D. R., Kamataki, T., Waxman, D. J., Guengerich, F. P., Estabrook, R. W., Feyereisen, R., Gonzalez, F. J., Coon, M. J., Gunsalus, I. C., Gotoh, O., Okuda, K. and Nebert, D. W. (1993) DNA Cell. Biol. **12**, 1–51
2 Nebert, D. W. and Gonzalez, F. J. (1987) Annu. Rev. Biochem. **56**, 945–993
3 Nelson, D. R. and Strobel, H. W. (1987) Mol. Biol. Evol. **4**, 572–593
4 Gonzalez, F. J. and Nebert, D. W. (1990) Trends Genet. **6**, 182–186
5 Gonzalez, F. J. (1989) Pharmacol. Rev. **40**, 243–287
6 Stegeman, J. J. and Hahn, M. E. (1994) in Aquatic Toxicology: Cellular, Molecular and Biochemical Perspectives (Malins, D. and Ostrander, G., eds.), pp. 87–206, Lewis Press, Boca Raton
7 Jaiswal, A. K., Gonzalez, F. J. and Nebert, D. W. (1985) Science **228**, 80–83
8 Conney, A. H. (1982) Cancer Res. **42**, 4875–4917
9 Ioannides, C. and Parke, D. V. (1990) Drug Metab. Rev. **22**, 1–85
10 Fagan, J. B., Pastewka, J. V., Chalberg, S. C., Gozukara, E., Guengerich, F. P. and Gelboin, H. V. (1986) Arch. Biochem. Biophys. **244**, 261–272
11 Strom, D. K., Postlind, H. and Tukey, R. H. (1992) Arch. Biochem. Biophys. **294**, 707–716
12 Quattrochi, L. C., Vu, T. and Tukey, R. H. (1994) J. Biol. Chem. **269**, 6949–6954
13 Heilmann, L. J., Sheen, Y.-Y., Bigelow, S. W. and Nebert, D. W. (1988) DNA **7**, 379–387
14 Berndtson, A. and Chen, T. T. (1994) Arch. Biochem. Biophys. **310**, 187–195
15 Stegeman, J. J., Smolowitz, R. M. and Hahn, M. E. (1991) Toxicol. Appl. Pharmacol. **110**, 486–504
16 Hahn, M. E. and Stegeman, J. J. (1994) Toxicol. Appl. Pharm. **127**, 187–198
17 Clemens, M. J. (1984) in Transcription and Translation: A Practical Approach (Hames, B. D. and Higgins, S. J., eds.), pp. 211–230, IRL Press, Oxford
18 Sambrook, J., Fritsch, E. F. and Maniatis, T. (eds.)(1984) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
19 Frohman, M. A., Downs, T. R., Chomczynski, P. and Frohman, L. A. (1989) Mol. Endocrinol. **3**, 1529–1536
20 Higgins, D. G. and Sharp, P. M. (1988) Gene **73**, 237–244
21 Olsen, G. J. (1988) Methods Enzymol. **164**, 793–812
22 Swofford, D. L. (1990) PAUP: Phylogenetic Analysis Using Parsimony, Illinois Natural History Survey, Champaign, IL
23 Klotz, A. V., Stegeman, J. J. and Walsh, C. (1983) Arch. Biochem. Biophys. **226**, 578–592
24 Graur, D., Hide, W. A., Zharkikh, A. and Li, W.-H. (1992) Comp. Biochem. Physiol. **101B**, 495–498
25 de Jong, W. W., Gleaves, J. T. and Boulter, D. (1977) J. Mol. Evol. **10**, 123–135
26 Shoshani, J. (1986) Mol. Biol. Evol. **3**, 222–242
27 Wyss, A. R., Novacek, M. J. and McKenna, M. C. (1987) Mol. Biol. Evol. **4**, 99–116
28 Easteal, S. (1990) Genetics **124**, 165–173
29 Sinclair, J. F. and Sinclair, P. R. (1993) in Cytochrome P450, Handbook of Experimental Toxicology (Schenkman, J. B. and Greim, H., eds.), pp. 259–277, Springer-Verlag, New York
30 Rifkind, A. B., Kanetoshi, A., Orlinick, J., Capdevila, J. H. and Lee, C. (1994) J. Biol. Chem. **269**, 3387–3396
31 Kloepper-Sams, P. J. (1989) Molecular Regulation of the Induction of Cytochrome P-450E in the Estuarine Fish *Fundulus heteroclitus*. Woods Hole Oceanographic Institution/Massachusetts Institute of Technology, Doctoral Dissertation
32 Leaver, M. J., Pirrit, L. and George, S. G. (1993) Mol. Marine Biol. Biotechnol. **2**, 338–345
33 Stegeman, J. J. (1989) Xenobiotica **19**, 1093–1110
34 Edwards, R. J., Singleton, A. M., Murray, B. P., Murray, S., Boobis, A. R. and Davies, D. S. (1991) Biochem. J. **278**, 749–757
35 Kloepper-Sams, P. J., Park, S. S., Gelboin, H. V. and Stegeman, J. J. (1987) Arch. Biochem. Biophys. **253**, 268–278
36 Sutter, T. R., Tang, Y. M., Hayes, C. L., Wo, Y.-Y. P., Jabs, E. W., Li, X., Yin, H., Cody, C. W. and Greenlee, W. F. (1994) J. Biol. Chem. **269**, 13092–13099
37 Hahn, M. E., Poland, A., Glover, E. and Stegeman, J. J. (1994) Arch. Biochem. Biophys. **310**, 218–228