Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 6731–6735, September 1986
Biochemistry

Human cytochrome P-450 4 mRNA and gene: Part of a multigene family that contains *Alu* sequences in its mRNA

(human repetitive sequences/intron–exon junctions)

LINDA C. QUATTROCHI, USHA R. PENDURTHI, STEVE T. OKINO, CAROL POTENZA, AND ROBERT H. TUKEY*

Department of Medicine, Cancer Center and Division of Pharmacology, T-012, University of California, San Diego, La Jolla, CA 92093

Communicated by Ronald W. Estabrook, May 27, 1986

ABSTRACT **Several overlapping λgt11 cDNA clones have been sequenced and shown to encode for the full-length human cytochrome P-450 4. The structure and location of the exons and flanking intron regions were also identified from a λEMBL-3 human genomic clone that encodes the full-length human P-450 4 gene. The human P-450 4 mRNA is flanked by 62 base pairs of 5′- and 1508 base pairs of 3′-noncoding sequence, with 1548 bases that encode a protein of 516 amino acids (M_r, 58,376). The predicted amino acid sequence of human P-450 4 is 69% and 70% homologous to its equivalent in mouse and rat, respectively, 75% homologous to rabbit P-450 4, and 68% homologous to human P_1-450. The 7.6-kilobase gene encodes 3118 nucleotides of exon sequence that is separated by six introns into seven exons. Exon 7, which is 1802 nucleotides, contains three inverse/complement *Alu* sequences that are organized in tandem. Comparison of the genomic DNA sequence of the human P-450 4 gene with the human P_1-450 and related genes in rat and mouse and the identification of the amino acid residues and triplet codon at each exon–intron junction show that the location of each intron in the human P-450 4 gene is conserved within this gene family. Although the length and homology of the introns within a related gene family may not be conserved, the location of intronic sequences may be an important determinant in the identification of related P-450 genes.**

The cytochrome P-450 monooxygenases are a group of proteins responsible for the oxidation of drugs, environmental pollutants, and endogenous compounds such as steroids and fatty acids (1). The diversity within this group of enzymes is dictated by structural heterogeneity, tissue specificity, and the differential regulation of the various forms of P-450. One highly conserved subfamily of the P-450 superfamily is represented by two forms of P-450 that are susceptible to induction by polycyclic aromatic hydrocarbons. In rabbits, cDNAs that encode for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (Cl_4Bz_2-dioxin)-inducible P-450 4 and P-450 6 have been shown by DNA sequence analysis to be homologous to rat P-450d and P-450c and mouse P_3-450 and P_1-450, respectively (2). Rabbit P-450 4 and P-450 6 are differentially regulated during development, with P-450 4 constituting the major adult form in liver while P-450 6 is expressed in other tissues (3, 4). Similar observations have been made for the homologous forms in the mouse (5). In mice and rats, it has been shown that the induction by Cl_4Bz_2-dioxin requires the presence of specific DNA regions on the 5′ side of the cap site on the P_1-450 and P-450c (6–8) genes.

An important step to understand how this gene family has evolved and what mechanisms control its expression in humans is to characterize the mRNAs and genes that encode this family of P-450s. The cDNA and gene that encode the human equivalent to mouse P_1-450 has been fully characterized (9, 10). It has been proposed that in humans, the gene equivalent to the major Cl_4Bz_2-dioxin-inducible form in rabbit, P-450 4, and its counterparts in mice and rats, P_3-450 and P-450d, respectively, does not exist (9). However, our laboratory has shown by DNA sequence analysis of a 600-base-pair (bp) cDNA clone called hpP-450 4, which was identified by hybridization of a rabbit P-450 4 cDNA clone (2) to a human cDNA library, that the human equivalent to rabbit P-450 4 gene is expressed in human liver (11). Southern blot analysis also showed that the P-450 4 and P-450 6 (P_1-450) mRNAs were encoded by separate genes. Support for the existence of the P-450 4 has also been shown by the identification of a human liver P-450 that is homologous to rat P-450d (12). In the present study, we have used clone hpP-450 4 to isolate and characterize the full-length human P-450 4 mRNA and its corresponding structural gene, concluding that the human liver contains both members of the Cl_4Bz_2-dioxin-inducible gene family.

MATERIALS AND METHODS

Materials. Restriction endonucleases were purchased from New England Biolabs. Other modifying enzymes used in the construction of the cDNA and genomic libraries were purchased from Pharmacia P-L Biochemicals. DNA polymerase used in the DNA sequencing reactions was purchased from Bethesda Research Laboratories. DNA sequencing reagents and labeled deoxynucleotides were purchased from Amersham. λgt11 and EMBL-3 cloning vectors, *Escherichia coli* 1088 and 1090 were obtained from Vector Cloning Systems (Sorrento Valley, CA). *E. coli* K802 was generously supplied by Frank Gonzalez (National Institutes of Health).

Isolation of Full-Length Human P-450 4 cDNA and Gene. We had previously isolated and characterized from a human pBR322 cDNA library a 600-bp clone, hpP-450 4, that was isolated by hybridization to a rabbit P-450 4 cDNA clone (11). To obtain clones that span the entire human P-450 4 mRNA, a human λgt11 cDNA library was constructed with mRNAs size-fractionated on a 5–35% (wt/vol) sucrose density gradient (13), and the library was screened with a nick-translated insert isolated from clone hpP-450 4. Clones that hybridized under stringent wash conditions were purified, and the cDNA inserts were characterized by restriction endonuclease analysis and blot hybridization analysis. Selected fragments from the inserts were subcloned for nucleotide sequencing.

Human liver DNA, isolated from the same sample of tissue used to construct the λgt11 cDNA library, was partially digested with *Mbo* I, and the DNA was size-fractionated on a 5–35% (wt/vol) sucrose density gradient. Fragments that exceeded 15 kilobases (kb) were ligated into the *Bam*HI site of the λ replacement vector EMBL-3 (14), packaged, and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Cl_4Bz_2-dioxin, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin; bp, base pairs; kb, kilobases.

*To whom reprint requests should be addressed.

plated onto *E. coli* K802; and the library was screened with the nick-translated insert derived from clone hpP-450 4. Two genomic clones were selected and determined to encode the full-length gene based on their similar restriction endonuclease mapping and selective hybridization with the 5′ and 3′ portions of the full-length cDNA clones. Exons and introns were initially identified by hybridization of selective nick-translated cDNA probes, and further defined by DNA sequence analysis originating from restriction endonuclease sites that were shown to be conserved in the cDNA clones. This approach resulted in DNA sequence analysis of all the exon regions and a portion of each respective intron.

Nucleotide Sequence Analysis. The nucleotide sequences of the cDNAs and gene were determined from selected fragments inserted into either M13mp10, M13mp11, or M13mp19 by the dideoxy chain-termination method (15, 16).

RNA Blot Analysis. RNA was isolated by the guanidinium·HCL method (17) followed by purification of the mRNA by repeated oligo(dT)-cellulose chromatography (18). RNA gel blot analysis (19) was performed as outlined (20).

Primer Extension Analysis. Location of the cap site of the human P-450 4 mRNA was determined by the method of primer extension analysis (21) using a 15-bp oligonucleotide directed against the sequence TACAGATGGCATTGT.

RESULTS AND DISCUSSION

Identification of the Human P-450 4 mRNA and Structural Gene. Using clone hpP-450 4 (11) as a probe, the human cDNA and genomic clones corresponding to human P-450 4 were isolated. DNA sequence of several overlapping λgt11 cDNA clones, as well as the respective exons, exon and intron junctions, and part of each intron from a genomic EMBL-3 clone, was determined. Fig. 1 shows the restriction endonuclease sites and strategy for sequencing the cDNA and genomic clones. Fig. 2 shows the DNA sequence of the full-length human P-450 mRNA determined from the complete sequence of both the cDNA and most of the corresponding exon regions from the gene. The DNA sequences of the cDNAs and genomic clone in the region that hybridized with clone hpP-450 4 were the same showing that clones homologous to hpP-450 4 had been isolated.

Characterization of Human P-450 4 cDNAs. The length of the mRNA derived from the four overlapping cDNA clones was 3101 nucleotides (Fig. 2). However, 5′-extension analysis, using a 15-bp oligonucleotide directed to anneal starting at base 57 of the human P-450 4 mRNA, extended the location of the cap site 17 bases (data not shown), indicating the actual size of the mRNA to be 3118 bases long. The human P-450 4 mRNA encodes a 62-bp 5′- and a 1508-bp 3′-noncoding region and 1548 bases that encode a protein of 516 amino acids based on the open reading frame that initiated from the ATG codon at base position 63. The 62-bp 5′-noncoding region is conserved in length and is 70% homologous to the same region in rat P-450d (22) and mouse P_3-450 mRNAs (23), with little homology toward rat P-450c (24) or mouse and human P_1-450 mRNAs (9, 23). The location of the ATG codon encoding the initiation methionine is supported by the perfect homology of the putative NH_3-terminal sequence of the human P-450 4 cDNA and the 18 NH_3-terminal amino acids determined from HLd (12), a human P-450 that was immunopurified from microsomes with antibodies specific for rat P-450d (25). Based upon the high degree of homology between the predicted NH_3-terminal amino acid sequence and HLd, we would predict that the human mRNA encodes the equivalent to rat P-450d (26), which has been shown to be conserved in DNA and amino acid sequence with mouse P_3-450 (23) and rabbit P-450 4 (2).

Among cytochrome P-450s, the predicted amino acid sequence of human P-450 4 is 68, 69, and 70% homologous to the predicted sequence from cDNA clones encoding the entire human P_1-450 (9), mouse P_3-450 (23), and rat P-450d (26), respectively. However, 78% homology was exhibited toward the predicted amino acid sequence from the rabbit P-450 4 cDNA clone, which encoded 83% of the protein (2). Two base pair changes in our previous sequence of the rabbit P-450 4 cDNA (2) changes leucine-356 to an arginine and isoleucine-461 to a threonine. The amino acid sequence of rabbit form 4 (27) has been deduced from the isolated protein and differs in nine residues from the cDNA sequence (2). Although there appears to be diversity in the reported structure of rabbit P-450 4 from several different sources (2, 27, 28), the complete rabbit P-450 4 sequence (27) is 75% homologous to human P-450 4. It is of interest to note that with a few exceptions, amino acids that confer structural integrity to the proteins, such as cysteine and proline residues, are highly conserved between human P-450 4 and its counterparts in rodents and rabbits. The degree of homology exhibited between human P-450 4 and human P_1-450 is comparable to the homology observed between the related forms in rats (24, 26), mice (23), and rabbit (2).

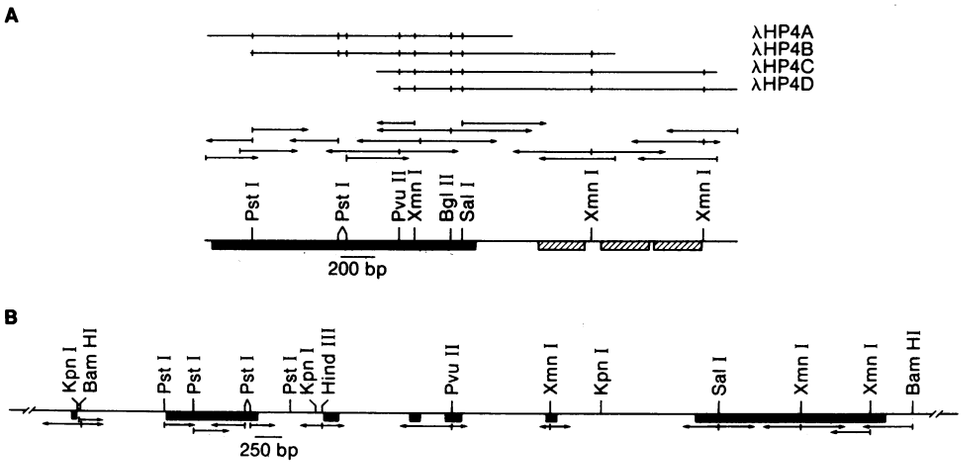

FIG. 1. Restriction endonuclease maps and sequence strategy for the human P-450 4 cDNA and genomic clones. (*A*) Four overlapping cDNA clones used to determine the complete cDNA sequence are shown. Arrows below the restriction maps indicate the direction of sequencing of each fragment. Although only one arrow is represented for each fragment, multiple fragments were usually generated from each subclone and subjected to DNA sequence analysis. (*B*) The restriction map and exon–intron organization of the human P-450 4 gene is shown. Solid boxes projecting below the line represent exons. Below the restriction map is the sequencing strategy for determining the exon–intron junctions.

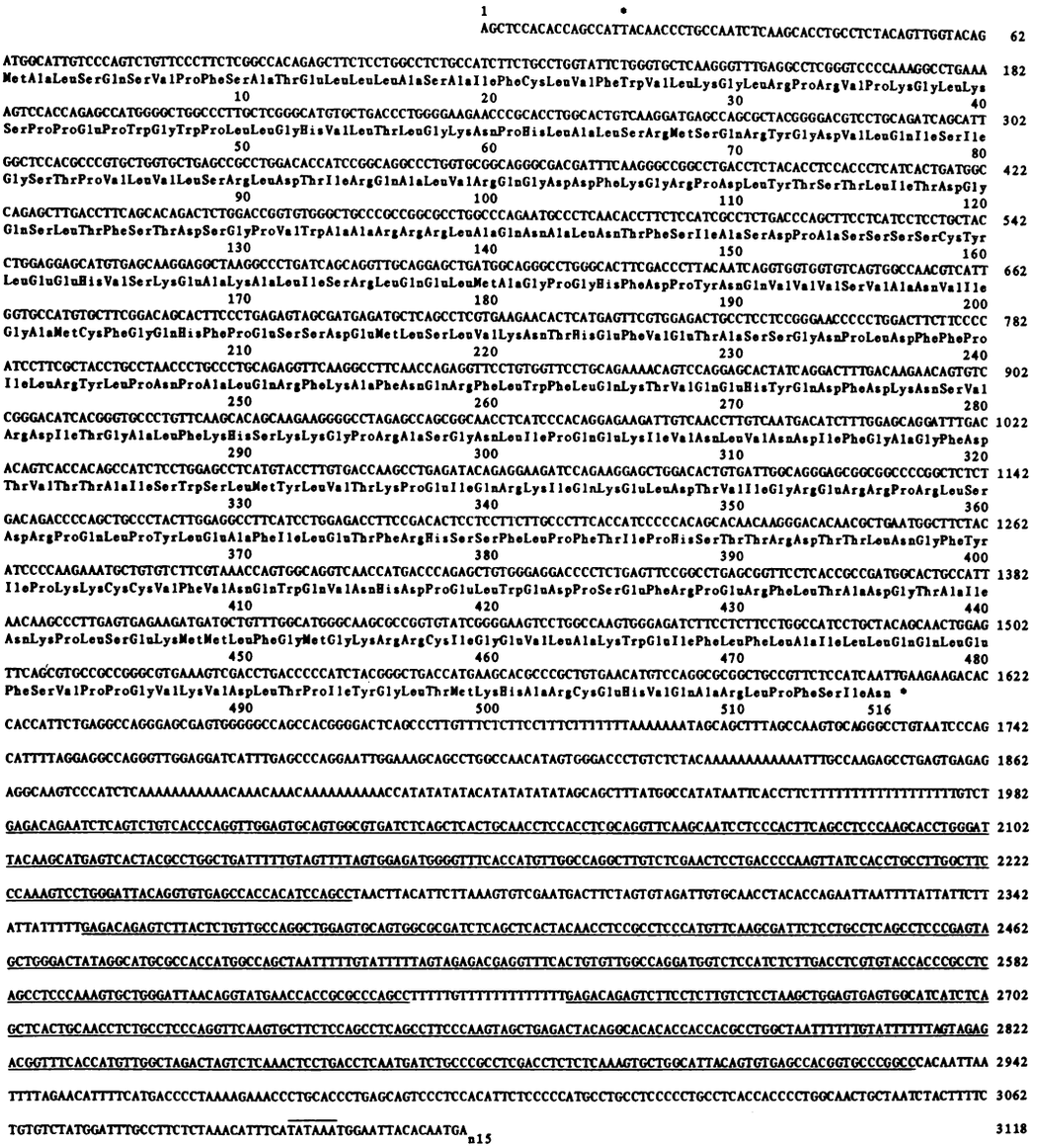

FIG. 2. Nucleotide sequence of the human P-450 4 mRNA. The complete sequence of the P-450 4 exonic regions, derived from the overlapping cDNA clones and the exon sequences of the gene, is shown. Along the side of the sequence are the numbers for the nucleotides. Nucleotide number 1 is the predicted site for the initiation of transcription as determined from 5′-extension analysis. The asterisk (*) above base 18 is the 5′ end of cDNA clone λHP4A (Fig. 1). The numbering of the predicted amino acids is indicated underneath the sequence. The three *Alu* sequences in the 3′-untranslation region are underlined. A putative polyadenylylation signal is overlined on the last line of the sequence.

The 3′-Untranslated Region of the mRNA Contains *Alu* Sequences. The 3′-noncoding region of human P-450 4, which spans 1508 nucleotides, completely diverges in sequence homology with the other homologous P-450s. The striking feature of this region, which contributes to its unusual length, is the presence of three repetitive sequences of the human *Alu* family (Fig. 2). A consensus sequence, derived by Deininger *et al.* (29), consists of DNA, approximately 300 bp long, flanked at the 3′ end with adenosine-rich sequences and bounded by direct repeats. *Alu* sequences constitute 3–6% the total mass of DNA (30) and are usually located in intergenic DNA or in introns. The only mature mRNAs noted that contain *Alu*-like sequences are those of the class I histocompatibility antigens of the mouse (31), one of the glutathione *S*-transferases (32), and the human low density lipoprotein receptor mRNA (33), which contains several repetitive *Alu* sequences in the 3′-noncoding region. The *Alu* sequences located in the human P-450 4 mRNA, which are 83–85% homologous to the consensus sequence (29), are organized in an inverse/complement fashion with a thymidine-rich region immediately to the 5′ side of each sequence, and are not flanked by direct repeats. Direct repeats are common to most *Alu* sequences, indicating that they integrate via a mechanism that has been proposed to be common to mobile genetic elements (34). Since direct repeats do not flank the *Alu* sequences in the human P-450 4 mRNA, it is difficult to predict the origin of these repetitive elements.

Although the appearance of *Alu* sequences in mature mRNA is not a common observation, these sequences appear to be present in other human mRNAs. Shown in Fig. 3 is an RNA gel blot of human, mouse, and rabbit mRNA probed with the 247-bp *Pst* I fragment from exon 2 and the human P-450 4 λHP4B cDNA (Fig. 1), a clone that contained the majority of the coding region and the first *Alu* sequence. As shown (11), exon sequences from human P-450 4 preferentially hybridize to Cl_4Bz_2-dioxin-inducible mouse P_3-450 and rabbit P-450 4, a result which is explained by the greater degree of homology of the human P-450 4 cDNA to those sequences. In addition, when the human liver mRNA is probed, the 3.1-kb P-450 4 mRNA is seen as well as faint

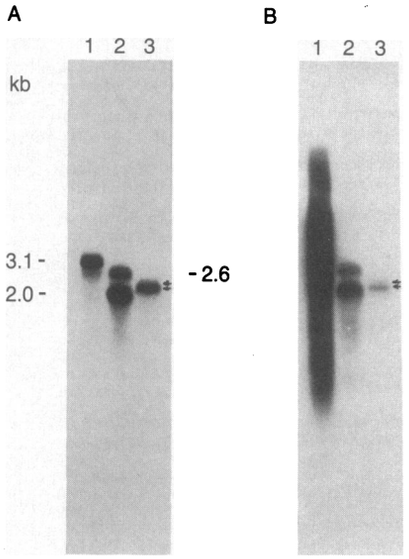

FIG. 3. RNA gel blot analysis of human, mouse, and rabbit mRNA probed with the human P-450 4 gene and cDNA. Preparations of liver mRNA from human (lanes 1), Cl_4Bz_2-dioxin-treated C57/B6 mice (lanes 2), and Cl_4Bz_2-dioxin-treated rabbit (lanes 3) were electrophoresed in duplicate in 0.8% agarose gels containing 10 mM methylmercury hydroxide and transferred to diazobenzyloxymethyl paper. (*A*) The 5′ 247-bp *Pst* I fragment, containing a portion of exon 2 of the human P-450 4 gene, was used as a probe to hybridize to the mRNAs. (*B*) Hybridization of the mRNAs to the coding and noncoding cDNA insert of hpP-450 4 is shown for the duplicate gel. The arrows mark the location of the Cl_4Bz_2-dioxin-inducible P-450 6 (upper arrow) and P-450 4 (lower arrow), while the 2.6-kb and 2.0-kb markers are aligned with the mouse P_1-450 and P_3-450 mRNAs, respectively. The conditions for treatment of the animals with Cl_4Bz_2-dioxin are as outlined (2).

hybridization to the 2.6-kb human P_1-450 mRNA. When clone HP4B was used as a probe, hybridization occurred to the mouse and rabbit mRNA. However, in the lane that contained the human mRNA, the HP4B probe caused a smearing pattern with no definition of the human P-450 4 mRNA. Since *Alu* elements are found only in human DNA, the smearing pattern seen in the lane with the human mRNA indicates the presence of *Alu* sequences in other mature mRNAs.

Characterization of Human P-450 4 Structural Gene. Since we had shown that clone hpP-450 4 was homologous to human P_1-450 mRNA (11), hybridization and washing conditions were chosen to minimize cross-hybridization to the human P_1-450 gene sequences. Two genomic clones approximately 12 kb long with similar restriction endonuclease sites were characterized and shown by Southern blot and hybridization analysis with 5′ and 3′ portions of the full-length cDNA clones to encode the full-length gene. Selected fragments were subcloned to facilitate DNA sequencing of most of the exon regions and all of the intron–exon junctions. The DNA sequences of the exon regions were identical to the corresponding cDNA sequence, validating the identity of the human P-450 4 gene.

The organization of the human P-450 4 gene is similar to the rat P-450c (35) and P-450d genes (22), the mouse P_3-450 and P_1-450 genes (36) and human P_1-450 gene (10) with seven exons and six introns (Fig. 4). The first exon, which is 52 bases long and lies 24 bp from the consensus TATAA (Fig. 4) box, is 70% homologous to the first exon of the rat P-450d (22) and mouse P_3-450 genes (36). All introns begin with the nucleotides G-T and end with the nucleotides A-G (Fig. 4), which is consistent with the consensus sequence for exon–intron splice junctions for eukaryotic genes (37). The lengths of the exons are 52, 840, 121, 90, 124, 87, and 1803 nucleotides in their 5′ to 3′ order. The insertion site of the six introns in the human P-450 4 gene is rigidly conserved between the rodent and human genes, with the exception of the exon–intron junction of exon 5 where the split codon is a proline in the human gene, a glutamic acid in the rodent genes and a glutamine in the human P_1-450 gene. Although the lengths of exons 2 and 3 differ from the mouse and rat by a few bases, the largest difference occurs in human exon 7, which is 1216 bases larger, part of which is the result of the insertion of three *Alu* sequences.

The approximated lengths of each intron starting from the first intron are 850, 590, 725, 266, 800, and 1330 nucleotides. With the exception of intron 4, which is 11 bases shorter than intron 4 in the mouse P_3-450 gene (36), we did not pursue completing the entire structure of all the intervening sequences of the human P-450 4 gene. Although exact lengths could not be established for all the introns, restriction endonuclease maps indicated that slight differences exist in the length of each intron between the human and rodent genes, with the greatest difference occurring in intron 1, estimated to be 200 bases shorter in the human P-450 4 gene. Combined, the predicted length of the human P-450 4 gene is

```
                                                                -1  1
5' Flanking  TATAGGGGCGTCTTTATAAAAAGGCCACTCACCTAGAGCCAGA  AGCTCCACACCA

                    52          INTRON
First exon  ...CTACAG GTACC..(1-850 bp)...TGCAG TTGGTA...Second exon

                   893
Second exon ...GACAAG GTGAG..(2-590 bp)...CTCAG AACAGT....Third exon
               AspLys                          Asn

                  1014
Third exon.....GAGCAG GTAGG..(3-725 bp)...TTCAG GATTTG...Fourth exon
                 AlaG                          lyPhe

                  1104
Fourth exon....AGCTGG GTACA..(4-266 bp)...TACAG ACACTG....Fifth exon
                 LeuA                          spThr

                  1228
Fifth exon.....CCACAG GTGAG..(5-800 bp)...CTCAG ÇACAAC....Sixth exon
                HisSe                          rThr

                  1315
Sixth exon.....TGACCC GTGAG..(6-1330 bp)..TGCAG AGAGCT..Seventh exon
                 AspPr                          oGlu
```

FIG. 4. Comparisons of the exon–intron junctions of the human P-450 4 gene. DNA sequence analysis of the intron–exon junctions and most of the corresponding exons was performed on specific fragments as indicated in Fig. 1. The 5′-flanking DNA, including the TATAA box (bold print) and several nucleotides of the first exon, are shown on the first line. In parentheses are the number and approximate length of each intron, flanked by the 5′ and 3′ nucleotides of each exon–intron junction. The consensus G-T and A-G bases that begin and end each intron are underlined. The location of each exon–intron junction is indicated by the nucleotide location and corresponding amino acids within the mRNA.

7650 nucleotides, larger than the 6.7-kb mouse P_3-450 gene (36), 6.9-kb rat P-450d gene (22), and 6.3-kb human P_1-450 gene (10).

Intron–Exon Organization and Comparison to Other P-450 Genes. The similarity in location of the introns in the human P-450 4 gene with the human P_1-450 gene (10), coupled with the high degree of homology displayed within the coding region and the similarity in length of the exons within this region, indicates that these genes are related and may have evolved from a common ancestral gene. The conservation in location of the introns is tightly controlled not only within species, but between different species.

The length and homology between the introns does not appear to be conserved to the same extent as between the exon sequences, which suggests that the evolutionary development of the genes may be influenced by common exonic requirements. Such would appear to be the case with the phenobarbital-inducible rat P-450b and P-450e genes (38), whose exonic sequences are separated by eight introns inserted in the same location in both genes. The organization of these genes is completely different than the intron–exon organization of the Cl_4Bz_2-dioxin-responsive genes. Although it has been proposed by Dayhoff *et al.* (39) that gene families may be related based on homology of the predicted amino acid sequence, it would appear that the similarity in location and number of introns is an important determinant in identifying related P-450 gene families. Since only a small number of P-450 genes have been characterized, a clear understanding of the evolutionary relationships between the P-450s may require the additional knowledge of the sequence and structure of other P-450 gene families.

The authors thank Dr. Steve Wrighton for providing us with a preprint of his manuscript (12) and Dr. Eric Johnson for helpful discussions. This work was supported in part by Grant CA37139 from the Public Health Service (R.H.T.). This work was conducted in part by the Clayton Foundation for Research California Division. R.H.T. is a Clayton Foundation Investigator.

1. Conney, A. H. (1982) *Cancer Res.* **42,** 4875–4917.
2. Okino, S. T., Quattrochi, L. C., Barnes, H. J., Osanto, S., Griffin, K. J., Johnson, E. F. & Tukey, R. H. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 5310–5314.
3. Norman, R. L., Johnson, E. F. & Muller-Eberhard, U. (1978) *J. Biol. Chem.* **253,** 8640–8647.
4. Liem, H. H., Muller-Eberhard, U. & Johnson, E. F. (1980) *Mol. Pharmacol.* **18,** 565–570.
5. Tuteja, N., Gonzalez, F. J. & Nebert, D. W. (1985) *Dev. Biol.* **112,** 117–184.
6. Jones, P. B. C., Galeazzi, D. R., Fisher, J. M. & Whitlock, J. P., Jr. (1985) *Science* **227,** 1499–1502.
7. Gonzalez, F. J. & Nebert, D. W. (1985) *Nucleic Acids Res.* **13,** 7269–7288.
8. Fujisawa-Sehara, A., Sogawa, K., Nishi, C. & Fujii-Kuriyama, Y. (1986) *Nucleic Acids Res.* **14,** 1465–1477.
9. Jaiswal, A. K., Gonzalez, F. J. & Nebert, D. W. (1985) *Science* **228,** 80–83.
10. Jaiswal, A. K., Gonzalez, F. J. & Nebert, D. W. (1985) *Nucleic Acids Res.* **13,** 4503–4520.
11. Quattrochi, L. C., Okino, S. T., Pendurthi, U. R. & Tukey, R. H. (1985) *DNA* **4,** 395–400.
12. Wrighton, S. A., Campanile, C., Thomas, P. E., Maines, S. L., Watkins, P. B., Parker, G., Mendez-Picon, G., Haniu, M., Shively, J. E., Levin, W. & Guzelian, P. S. (1986) *Mol. Pharmacol.* **29,** 405–410.
13. Young, R. & Davis, R. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1194–1198.
14. Frischauf, A-M., Lehrach, H., Poustka, A. & Murray, N. (1983) *J. Mol. Biol.* **170,** 827–842.
15. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143,** 161–178.
16. Messing, J. & Vieira, J. (1982) *Gene* **19,** 269–276.
17. Cox, R. A. (1968) *Methods Enzymol.* **12,** 120–129.
18. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
19. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5350–5354.
20. Tukey, R. H., Nebert, D. W. & Negishi, M. (1981) *J. Biol. Chem.* **256,** 6969–6974.
21. Hagenbuchle, O., Bovery, R. & Young, R. A. (1980) *Cell* **21,** 179–187.
22. Sogawa, K., Gotoh, O., Kawajiri, K., Harada, T. & Fujii-Kuriyama, Y. (1985) *J. Biol. Chem.* **260,** 5026–5032.
23. Kimura, S., Gonzalez, F. J. & Nebert, D. W. (1984) *J. Biol. Chem.* **259,** 10705–10713.
24. Yabusaki, Y., Shimizu, M., Murakami, H., Nakamura, K., Oeda, K. & Ohkawa, H. (1984) *Nucleic Acids Res.* **12,** 2929–2938.
25. Reik, L. M., Levin, W., Ryan, D. E. & Thomas, P. E. (1982) *J. Biol. Chem.* **257,** 3950–3957.
26. Kawajiri, K., Gotoh, O., Sogawa, K., Tagashira, Y., Muramatsu, M. & Fujii-Kuriyama, Y. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1649–1653.
27. Ozols, J. (1986) *J. Biol. Chem.* **261,** 3965–3979.
28. Fujita, V. S., Black, S. D., Tarr, G. E., Koop, D. R. & Coon, M. J. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4260–4264.
29. Deininger, P. L., Jolly, D. J., Rubin, C. M., Friedman, T. & Schmid, C. (1981) *J. Mol. Biol.* **151,** 17–33.
30. Schmid, C. W. & Jelinek, W. R. (1982) *Science* **216,** 1065–1070.
31. Hood, L., Steinmetz, M. & Malissen, B. (1983) *Annu. Rev. Immunol.* **1,** 529–568.
32. Rothkopf, G. S., Telakowski-Hopkins, C. A., Stotish, R. L. & Pickett, C. B. (1986) *Biochemistry* **25,** 993–1002.
33. Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L. & Russell, D. W. (1984) *Cell* **39,** 27–38.
34. Sharp, P. A. (1983) *Nature (London)* **301,** 471–472.
35. Sogawa, K., Gotoh, O., Kawajiri, K. & Fujii-Kuriyama, Y. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 5066–5070.
36. Gonzalez, F. J., Kimura, S. & Nebert, D. W. (1985) *J. Biol. Chem.* **260,** 5040–5049.
37. Breathnech, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 4853–4857.
38. Suwa, Y., Mizukami, Y., Sogawa, K. & Fujii-Kuriyama, Y. (1985) *J. Biol. Chem.* **260,** 7980–7984.
39. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983) *Methods Enzymol.* **91,** 524–545.