*Proc. Natl. Acad. Sci. USA*
Vol. 83, pp. 2841–2845, May 1986
Biochemistry

# Complete nucleotide sequence of two steroid 21-hydroxylase genes tandemly arranged in human chromosome: A pseudogene and a genuine gene

(monooxygenase/steroid hormone/adrenal hyperplasia/gene cloning/molecular evolution)

YUJIRO HIGASHI*, HIDEFUMI YOSHIOKA*, MIYUKI YAMANE*, OSAMU GOTOH†, AND YOSHIAKI FUJII-KURIYAMA*

*Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, 1-37-1 Kami-Ikebukuro, Toshima-ku, Tokyo, 170 Japan; and †Department of Biochemistry, Saitama Cancer Center Research Institute, Ina-machi, Saitama, 362 Japan

*Communicated by I. C. Gunsalus, December 23, 1985*

ABSTRACT **Two 21-hydroxylase [*P-450(C21)*] genes have been isolated from a human genomic library using a bovine P-450(C21) cDNA. The insert DNAs containing the *P-450(C21)* genes were also hybridized with the sequences of the 5′ or 3′ end regions of human C4 cDNA, indicating a close linkage of the *P-450(C21)* gene to the *C4* gene. Sequence analysis has revealed that the two *P-450(C21)* genes are both ≈3.4 kilobases long and split into 10 exons. Comparing the two sequences, we found that the two genes are highly homologous including their introns and flanking sequences, but that three mutations render one of the two *P-450(C21)* genes nonfunctional—1 base insertion, an 8-base deletion, and a transition mutation—all of which may cause premature termination of the translation. Tandem arrangement of the highly homologous pseudo- and genuine genes in close proximity could account for the high incidence of *P-450(C21)* gene deficiency by homologous gene recombination.**

The adrenal steroid 21-hydroxylase [*P-450(C21)*] gene belongs to the cytochrome P-450 supergene family and plays a crucial role in the synthesis of steroid hormones such as cortisol and aldosterone (1). In humans, congenital adrenal hyperplasia is observed in a rather high frequency (≈1 in 5000 births) and, thus, is one of the most common inborn errors of metabolism. This disease results from a deficiency in one of the enzymes involved in steroidogenesis. Approximately 95% of the affected cases have been reported to be due to a defect only in the P-450(C21) enzyme with an autosomal recessive trait closely linked to the HLA major histocompatibility complex (2).

Recent studies using gene cloning techniques have shown that there are two *P-450(C21)* genes, each located near the 3′ end of one of the two *C4* genes in a relatively short stretch of human chromosomal DNA (3, 4). White *et al.* have found that deletion of one of the two *P-450(C21)* genes was closely associated with the P-450(C21) deficiency in the patients with *HLA-Bw47* haplotype, raising the possibility that the other gene for P-450(C21) may be nonfunctional (5). Motivated by interest in a high incidence of this genetic disease for the two responsible genes in human chromosome as well as in an evolutionary process of diversification of P-450 supergene family, we have cloned the two human *P-450(C21)* genes and determined their complete nucleotide sequences. In the present paper, we describe that the two genes for P-450(C21) are highly homologous in nucleotide sequence, but that three critical mutations were found in one of the two genes to render the gene nonfunctional. An implication of these observations in this genetic disorder and evolutionary aspect of the *P-450(C21)* gene are discussed.

## MATERIALS AND METHODS

**DNA Probes.** Bovine P-450(C21) cDNAs, pcP-450C21-1 and -2, containing the entire coding sequence for bovine P-450(C21) (unpublished results) were $^{32}$P-labeled by nick-translation ($\geqq 10^8$ cpm/$\mu$g) for the hybridization probes.

For the examination of the linkage between the *P-450(C21)* and *C4* genes, the two oligonucleotides, 5′ GAACAAGAGCAACCTGGGCT 3′ and 5′ CTGGCACCCCTGAGTGCCAT 3′, which are complementary to the 5′- and 3′-terminal coding sequences of the human C4 cDNA (6), respectively, were chemically synthesized and $^{32}$P-labeled ($2 \times 10^7$ cpm/$\mu$g) for the hybridization probes.

**Isolation of Genomic Clone for the *P-450(C21)* Gene.** A human genomic library cloned in Charon 4A (7) was screened with the bovine cDNA probes as described (8).

**DNA Preparation and Blot Hybridization Experiments.** Purification of the plasmid and phage DNAs (9) and blot hybridization analyses (10) were performed as described.

**DNA Sequence Analysis.** DNA sequencing was performed by the chain-termination method (11).

## RESULTS AND DISCUSSION

**Cloning of the *P-450(C21)* Gene and Its Linkage to the *C4* Gene.** A human genomic library (7) was screened with the cloned cDNAs containing the entire coding sequence for bovine P-450(C21). At least five of six different clones obtained from $10^6$ recombinant phages were finally classified into two independent groups by restriction mapping analysis. The remaining one (λC21C-1) seemed distinct from the two groups but not yet fully characterized. The restriction cleavage maps of the two groups are shown in Fig. 1*i*. Inserts of the two recombinant phages, λC21A-1 and λC21A-2, overlapping in part with each other, cover the sequence from 9 kilobases (kb) upstream to 12 kb downstream of one *P-450(C21)* gene. On the other hand, λC21B-1 and λC21B-2 also carried another *P-450(C21)* gene accompanied by the 9-kb upstream and 4-kb downstream flanking sequence.

Previous genetic studies of P-450(C21) deficiency have suggested a linkage of the *P-450(C21)* gene to the HLA complex region (2). Recently, close linkage between the human complement *C4* and *P-450(C21)* genes has been demonstrated by isolating several cosmid clones containing both *C4* and *P-450(C21)* genes in their single inserts (3, 4). To examine whether the isolated human *P-450(C21)* genes were linked to the *C4* genes, Southern blot analyses were performed with the cloned DNAs of the two groups by using the synthetic oligonucleotides as probes.

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: kb, kilobase(s); bp, base pair(s).

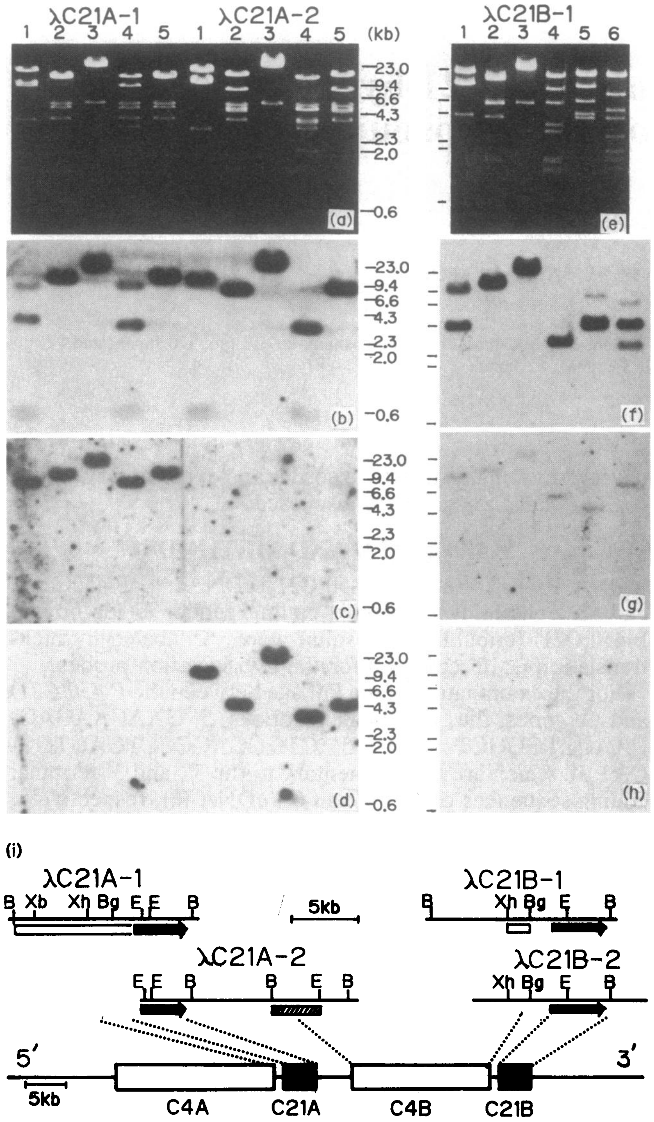

FIG. 1. Close linkage between *P-450(C21)* and *C4* genes. Ethidium bromide staining of the recombinant DNAs digested with various restriction enzymes in 0.8% agarose gel after electrophoresis: (*a*) λC21A-1 and λC21A-2; (*e*) λC21B-1. (*a*) Lanes 1, *Eco*RI; lanes 2, *Bam*HI; lanes 3, *Hin*dIII; lanes 4, *Eco*RI and *Bam*HI; lanes 5, *Bam*HI and *Hin*dIII. (*e*) Lane 1, *Eco*RI; lane 2, *Bam*HI; lane 3, *Hin*dIII; lane 4, *Kpn* I; lane 5, *Eco*RI and *Xho* I; lane 6, *Eco*RI and *Bgl* II. The blotted DNAs from the respective gels described above were hybridized with various $^{32}$P-labeled probe DNAs: (*b* and *f*) bovine P-450(C21) cDNA probe; (*c* and *g*) 3'-terminal oligonucleotide probe of human C4 cDNA; (*d* and *h*) 5'-terminal probe of human C4 cDNA. Numbers in between panels indicate the length of the marker DNA fragments (DNA digested with *Hin*dIII) in kb. (*i*) Restriction enzyme cleavage maps for λC21A-1 and -2, and λC21B-1 and -2. Bold arrows indicate the location and direction of the *P-450(C21)* genes. Hatched and open boxes show the hybridizing portions with the human C4 5'- and 3'-terminal oligonucleotide probes, respectively. The schematic representation for the tandemly arranged *C4* and *P-450(C21)* genes was taken from published data (5) and is shown at the bottom. Restriction enzyme cleavage sites are indicated as follows: E, *Eco*RI; B, *Bam*HI; Bg, *Bgl* II; Xh, *Xho* I; Xb, *Xba* I. The DNA transferred to filters was hybridized in 1 M NaCl/10 mM EDTA/0.1% NaDodSO$_4$/50 mM Tris·HCl, pH 7.5, at 48°C with the 5'-terminal probe or at 52°C for the 3'-terminal probe, and then washed in 0.9 M NaCl/0.09 M Na citrate at the same respective temperatures. Restriction maps were determined from restriction enzyme digestion experiments with the purified phage DNA and Southern blot hybridization analysis using a bovine P-450(C21) cDNA probe.

As shown in Fig. 1 *a–d*, λC21A-1 hybridized with the C4 3'-terminal probe, but not with the C4 5'-terminal probe, while λC21A-2 hybridized only with the C4 5'-terminal probe. As for the second recombinant phage DNA group, λC21B-1 also hybridized only with the C4 3'-terminal probe as shown in Fig. 1 *e–h*. These results, taken together with the restriction mappings of the cloned DNAs, clearly show that the two *P-450(C21)* genes are closely associated with the *C4* genes, and at least one *P-450(C21)* gene, which has been isolated in λC21A-1, seems to be sandwiched between the two *C4* genes. From comparison of the restriction maps of our phage clones with those of the cosmid clones reported recently (5), one *P-450(C21)* gene isolated in clone λC21A-1 corresponds to the *P-450(C21) "A"* gene and the second one in λC21B-1 is equivalent to the *P-450(C21) "B"* gene. The characteristic 3.2-kb and 3.7-kb *Taq* I fragments (5) are present in clone λC21A-1 and λC21B-1, respectively (see Fig. 3). These two *P-450(C21)* genes alternate with two *C4* genes in a relatively short stretch of chromosomal DNA as shown in Fig. 1*i*.

**Determination of the Complete Nucleotide Sequences of the *P-450(C21)* Genes.** We determined the complete nucleotide sequences of *P-450(C21) A* and *B* genes, which are shown in Fig. 3. The strategy for sequencing and the framework of exon and intron are shown only for the *B* gene in Fig. 2. Those for the *A* gene are essentially similar. The sequences for exons and introns were assigned with the aid of homology to the bovine P-450(C21) cDNA sequence (unpublished observations) and the GT–AG rule for exon–intron junction. All the exon–intron junctions follow the canonical GT–AG rule in the *P-450(C21) A* and *B* genes. The coding nucleotide sequence and the deduced amino acid sequence are ≈83% and ≈77% homologous, respectively, between the human gene and the bovine cDNA.

S1 nuclease mapping analysis showed three protected bands of different sizes with several faint bands (data not shown), each corresponding to the transcription start site at the −9th, −53rd, and −118th base position as shown in Fig. 3. Judging from the intensity of protected bands, the −9th position is a major transcription start site and is located 25 base pairs (bp) downstream from TATAA sequence at the −38th to −34th position. The two others at the −53rd and −118th positions are further upstream from the TATAA sequence and are presumably minor start sites because of lower intensity of the bands. Upstream from these positions, however, there is no apparent TATA sequence or its equiv-

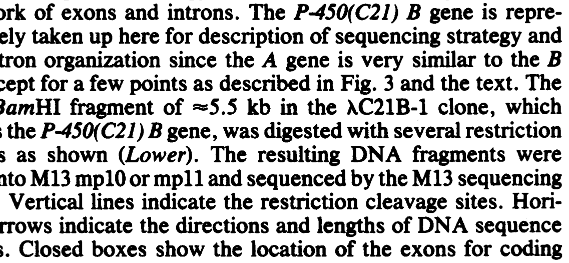

FIG. 2. Sequence strategy for human *P-450(C21)* gene and its framework of exons and introns. The *P-450(C21) B* gene is representatively taken up here for description of sequencing strategy and exon–intron organization since the *A* gene is very similar to the *B* gene except for a few points as described in Fig. 3 and the text. The *Bgl* II/*Bam*HI fragment of ≈5.5 kb in the λC21B-1 clone, which contains the *P-450(C21) B* gene, was digested with several restriction enzymes as shown (*Lower*). The resulting DNA fragments were cloned into M13 mp10 or mp11 and sequenced by the M13 sequencing method. Vertical lines indicate the restriction cleavage sites. Horizontal arrows indicate the directions and lengths of DNA sequence analyses. Closed boxes show the location of the exons for coding sequence. Exons are numbered from 1 to 10. B, *Bam*HI; Bg, *Bgl* II; E, *Eco*RI; Xh, *Xho* I. The scale is shown above the gene structure.

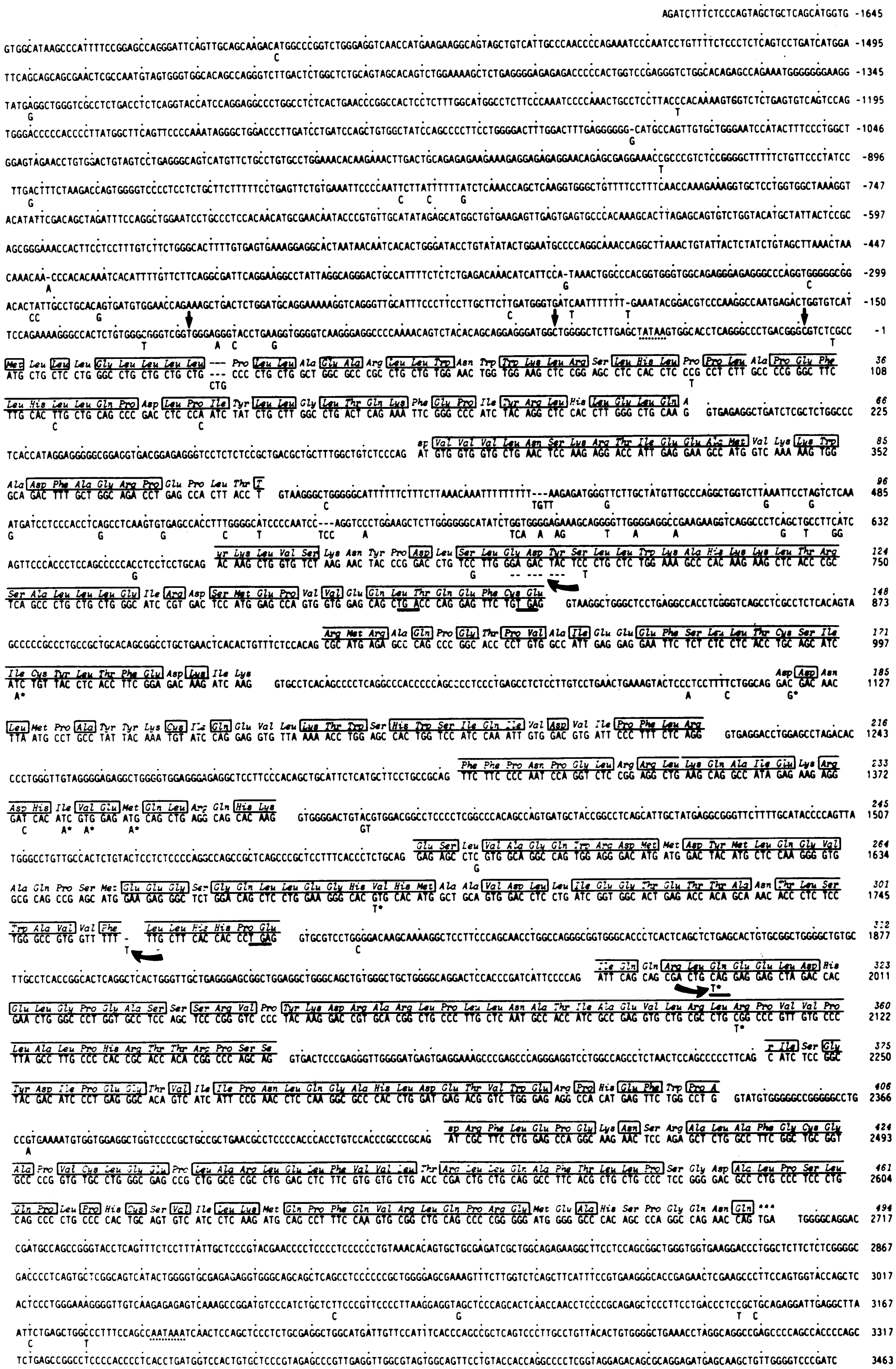

FIG. 3. Complete nucleotide sequence and deduced primary structure for the two human *P-450(C21)* genes. The complete nucleotide sequence and the deduced primary structure for the *P-450(C21) B* gene, an intact gene, are shown. The sequence is shown from the upstream *Bgl* II site to the downstream *Sau*3A site nearest to the *Bam*HI site (Fig. 2). Only nucleotide alterations in the *A* gene, a pseudogene, from the *B* counterpart are indicated under the corresponding nucleotides in the *B* gene. Of these alterations, base substitutions causing putative amino acid changes are marked by asterisks. Gaps, represented by bars, are introduced to minimize the differences presumably resulting from deletions or insertions in the nucleotide sequence. The three deleterious mutations in the *A* gene appeared as a deletion of 8 bp, GAGACTAC in the third exon, an insertion of 1 base (T) in the seventh exon, and a transition mutation of C to T in the eighth exon. They are indicated by upward arrows. In-phase termination codons resulting from frameshifts caused by such deleterious mutations are underlined. The transcription initiation sites determined by the S1 nuclease mapping analysis are indicated by downward arrows. The typical TATAA sequence and the poly(A) addition signal (AATAAA) are indicated by dotted lines. Amino acids identical for both human and bovine sequences are also enclosed in boxes.

alent for the transcription start site, analogous to the case with HMG CoA reductase gene (12).

At the 486 bp downstream from the termination codon in the two genes, there exists a typical poly(A) addition signal, AATAAA. By analogy with bovine P-450(C21) cDNA sequence, this signal may function as such in the case of humans. On the whole, the human *P-450(C21)* gene seems to be ≈3.4 kb long and contains 10 split exon sequences with very short 9 introns.

**Structural Comparison of the Two *P-450(C21)* Genes.** In the sequence of the *P-450(C21) A* gene (Fig. 3), we found several critical nucleotide alterations from the bovine cDNA sequence, which could not be accounted for by the difference in species. These alterations in the *A* gene are (*i*) an 8-bp deletion in the third exon, (*ii*) a 1-bp insertion in the seventh exon, and (*iii*) a transition (C–T) point mutation in the eighth exon (Fig. 3, arrows). All these presumably render the gene nonfunctional by generating premature terminations. On the other hand, the *B* gene has no such nucleotide alterations. In this gene, the 10 exon sequences assigned as described in the previous section provide an open reading frame for 494 amino acids, 2 amino acids less than the bovine counterpart. From these observations, we concluded that the *P-450(C21) A* gene is a pseudogene, whereas the *B* gene is a genuine gene. This conclusion was indeed substantiated by the RNA blot analysis of an adrenal total RNA preparation using specific probes for *A* and *B* gene sequences (sequences of the 701st to 720th and 695th to 723rd with a deletion of 8 bp and 2 base replacements for the B and A probes, respectively). Only a probe specific for the *B* gene yielded a hybridization band at ≈2.4 kb with the adrenal RNA preparation from a hormonally normal individual (unpublished data).

These results provide the structural basis for the observations reported by White *et al.* (5). From the DNA blot experiments of the *Taq* I-digested genomic DNAs derived from the patients and hormonally normal individuals, they suggested that the *P-450(C21) B* gene is functional, while the *A* gene is not.

Except for these critical base alterations, these two *P-450(C21)* genes show very extensive sequence homology with each other even in their flanking and intron sequences. In the region of ≈5.1 kb spanning from 1.5 kb upstream to 0.2 kb downstream of the *P-450(C21)* gene, only 88 base alterations were observed between the *A* and *B* genes, with an overall sequence homology of 98% (Fig. 3).

Recently, White *et al.* have reported the presence of two *P-450(C21)* genes in the mouse genome, each located immediately 3′ to the *C4* and *Slp* genes in the H-2 complex region (13), and the same situation has also been suggested in the bovine genome (14). It is reasonable, therefore, to infer that the close association of the *C4* and *P-450(C21)* genes and their duplication as a unit had occurred before the adaptive radiation of mammals (≈7 × $10^7$ years ago). The extensive homology observed between human *P-450(C21) A* and *B* genes could not be expected to have been maintained for such an evolutionary time period involved if these genes evolved independently under natural evolutionary pressure.

Independent evolution of the two intraspecies *P-450(C21)* genes would generate ≈20% sequence divergence between them as observed interspecifically between human and bovine coding sequences, although another possibility cannot be rigorously eliminated—recent and independent duplication of a set of *P-450(C21)* and *C4* genes in human, murine, and probably bovine genomes after the mammalian radiation. It seems highly probable that sequence divergence between

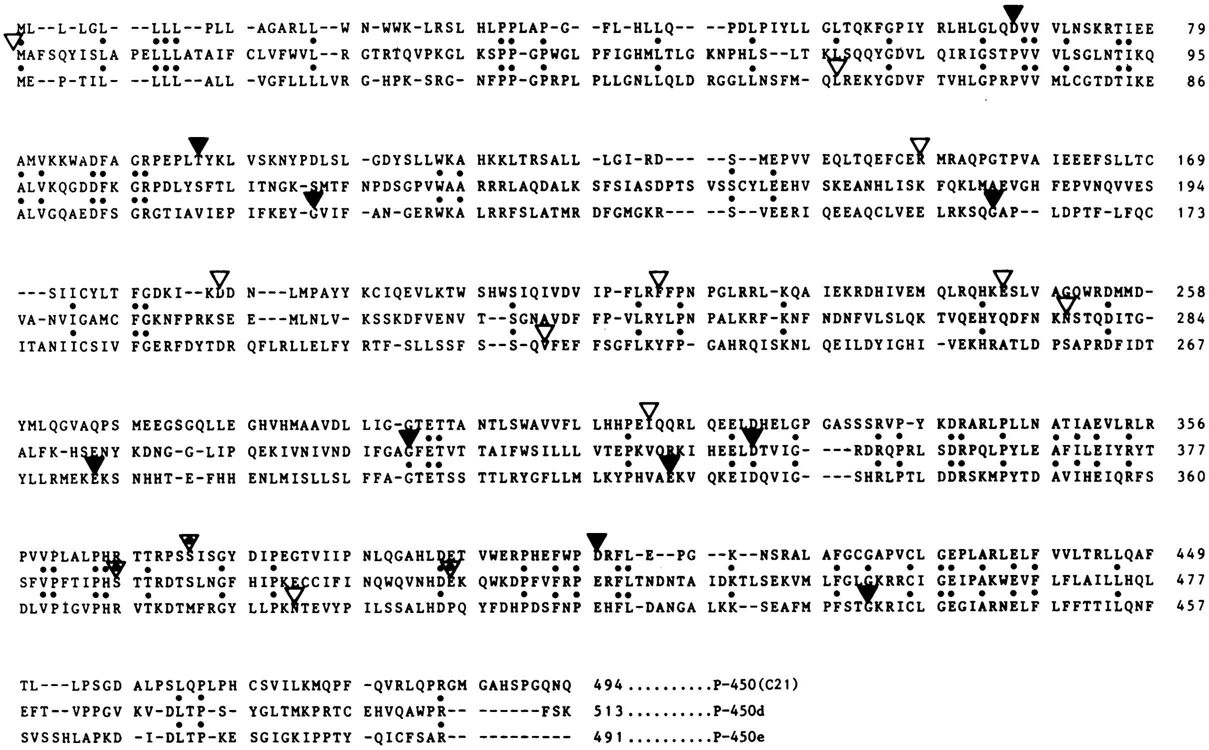
ML--L-LGL- --LLL--PLL -AGARLL--W N-WWK-LRSL HLPPLAP-G- -FL-HLLQ-- --PDLPIYLL GLTQKFGPIY RLHLGLQDVV VLNSKRTIEE 79
MAFSQYISLA PELLLATAIF CLVFWVL--R GTRTQVPKGL KSPP-GPWGL PFIGHMLTLG KNPHLS--LT KLSQQYGDVL QIRIGSTPVV VLSGLNTIKQ 95
ME--P-TIL- --LLL--ALL -VGFLLLLVR G-HPK-SRG- NFPP-GPRPL PLLGNLLQLD RGGLLNSFM- QLREKYGDVF TVHLGPRPVV MLCGTDTIKE 86
AMVKKWADFA GRPEPLTYKL VSKNYPDLSL -GDYSLLWKA HKKLTRSALL -LGI-RD--- --S--MEPVV EQLTQEFCER MRAQPGTPVA IEEEFSLLTC 169
ALVKQGDDFK GRPDLYSFTL ITNGK-SMTF NPDSGPVWAA RRRLAQDALK SFSIASDPTS VSSCYLEEHV SKEANHLISK FQKLMAEVGH FEPVNQVVES 194
ALVGQAEDFS GRGTIAVIEP IFKEY-GVIF -AN-GERWKA LRRFSLATMR DFGMGKR--- --S--VEERI QEEAQCLVEE LRKSQGAP-- LDPTF-LFQC 173
---SIICYLT FGDKI--KDD N---LMPAYY KCIQEVLKTW SHWSIQIVDV IP-FLRFFPN PGLRRL-KQA IEKRDHIVEM QLRQHKESLV AGQWRDMMD- 258
VA-NVIGAMC FGKNFPRKSE E---MLNLV- KSSKDFVENV T--SGNAVDF FP-VLRYLPN PALKRF-KNF NDNFVLSLQK TVQEHYQDFN KNSTQDITG- 284
ITANIICSIV FGERFDYTDR QFLRLLELFY RTF-SLLSSF S--S-QVFEF FSGFLKYFP- GAHRQISKNL QEILDYIGHI -VEKHRATLD PSAPRDFIDT 267
YMLQGVAQPS MEEGSGQLLE GHVHMAAVDL LIG-GTETTA NTLSWAVVFL LHHPEIQQRL QEELDHELGP GASSSRVP-Y KDRARLPLLN ATIAEVLRLR 356
ALFK-HSENY KDNG-G-LIP QEKIVNIVND IFGAGFETVT TAIFWSILLL VTEPKVQRKI HEELDTVIG- ---RDRQPRL SDRPQLPYLE AFILEIYRYT 377
YLLRMEKEKS NHHT-E-FHH ENLMISLLSL FFA-GTETSS TTLRYGFLLM LKYPHVAEKV QKEIDQVIG- ---SHRLPTL DDRSKMPYTD AVIHEIQRFS 360
PVVPLALPHR TTRPSSISGY DIPEGTVIIP NLQGAHLDET VWERPHEFWP DRFL-E--PG --K--NSRAL AFGCGAPVCL GEPLARLELF VVLTRLLQAF 449
SFVPFTIPHS TTRDTSLNGF HIPKECCIFI NQWQVNHDEK QWKDPFVFRP ERFLTNDNTA IDKTLSEKVM LFGLGKRRCI GEIPAKWEVF LFLAILLHQL 477
DLVPIGVPHR VTKDTMFRGY LLPKNTEVYP ILSSALHDPQ YFDHPDSFNP EHFL-DANGA LKK--SEAFM PFSTGKRICL GEGIARNELF LFFTTILQNF 457
TL---LPSGD ALPSLQPLPH CSVILKMQPF -QVRLQPRGM GAHSPGQNQ 494..........P-450(C21)
EFT--VPPGV KV-DLTP-S- YGLTMKPRTC EHVQAWPR-- ------FSK 513..........P-450d
SVSSHLAPKD -I-DLTP-KE SGIGKIPPTY -QICFSAR-- --------- 491..........P-450e

FIG. 4. Locations of introns in various forms of cytochrome P-450 genes in relation to the primary structures. Primary structures of P-450(C21), methylcholanthrene-inducible P-450d and phenobarbital-inducible P-450e, are represented by single-letter symbols. Gaps are introduced to maximize homology among three species of P-450s. Sites of introns are indicated in the amino acids involved as follows: ▿, introns localized immediately before the first nucleotide of the codon for the marked amino acids; ▾, introns localized between the first and second nucleotide of the codons for the marked amino acids; ▾, introns localized between the second and third nucleotide of the codons for the marked amino acids. The first intron of the *P-450d* gene is localized 14 bp upstream from the initiation (18) and its location is indicated by ▿ in front of the initiation methionine. Amino acids common to all three P-450s are indicated by dotting between them.

the two human genes has been rectified during this period of time by a mechanism such as concerted evolution involving gene conversion and/or unequal crossing-over, which has been actually observed in class I genes of the major histocompatibility complex (15) or in the human fetal γ-globin gene (16). An extensive sequence homology was also reported between the *C4* and *Slp* (*C4* equivalent) genes in the mouse genome (17). The presence of the duplicated and highly homologous sets of the *C4* and *P-450(C21)* genes in a short stretch of chromosomal DNA may allow for frequent exchange of their DNA sequences by homologous recombination or unequal crossing-over during meiosis. It is noticeable, from this viewpoint, that nearly half of the observed base changes between the *A* and *B* genes are localized in the second intron and the contiguous 5′ part of the third exon (Fig. 3). The presence of a hot spot in this limited region is reminiscent of a recombination point as suggested with the human fetal γ-globin gene (16).

In these genetic situations, once one of the two *P-450(C21)* genes became a pseudogene, the gene conversion or unequal crossing-over might have generated, more often than not, various gene recombination products unfavorable to the host animals. Some of them might have been fixed in the human genome as the affected allelic variant forms of the *P-450(C21)* gene. Such molecular events presumably occurring in the *P-450(C21)* gene could in part explain why, of the deficiencies in various steroidogenic P-450 enzymes, the *P-450(C21)* deficiency is the predominant cause of congenital adrenal hyperplasia (2). Detailed analysis of DNA from *P-450(C21)*-deficient patients will be required for precise understanding of the nature of this genetic disease.

**Exon–Intron Organization in P-450 Genes.** So far, gene structures for two types of P-450, phenobarbital- and methylcholanthrene-inducible P-450s (*P-450b* and *-e* and *P-450c* and *-d*, respectively), have been reported (18–23). In spite of the fact that they are supposed from their sequence homology to be originated by divergent evolution from a common ancestor, their exon–intron organizations are grossly different from each other (21). In Fig. 4, the locations of nine introns in the human *P-450(C21)* gene are represented in relation to the amino acid sequence, together with those in the two drug-inducible *P-450e* and *-d* genes for comparison. It is noticeable that no introns in the *P-450(C21)* genes find their exact counterpart in either of the two other P-450 genes with regard to the relative position of intron insertion. If possible shifts of intron locations due to deletion or insertion of coding sequences as reported for α-fetoprotein and albumin genes (24) are taken into consideration, it appears that the *P-450(C21)* gene shares two sites of introns with each of the two genes [the second and fifth introns in the *P-450(C21)* gene corresponding to the second and fourth ones in the *P-450e* gene, respectively, and the sixth and eighth introns in the *P-450(C21)* gene to the second and fourth ones in the *P-450d* gene, respectively]. Extensive alterations (deletion, insertion, and replacement) in amino acid sequences of the various P-450 molecules are likely to obscure the correspondence of the intron locations among these three P-450 genes. The finding that the *P-450(C21)* gene has possible common sites for introns with either of the two drug-inducible P-450 genes is consistent with the notion that the *P-450(C21)* gene is supposed to have separated from a common ancestor gene before the generation of the two drug-inducible P-450 genes. Although the convergent evolution of these P-450 genes from different ancestors cannot be rigorously eliminated, the other introns specific for each of the three genes might have been generated by either their independent insertion or deletion after the divergence of these genes from the ancestor gene (25). Elucidation of the gene structures for other types of P-450 may help us to understand the origin of the intron and the evolutionary process of diversification in the P-450 genes.

We thank Dr. T. Maniatis for the gift of the human genomic library and Dr. K. Sogawa for valuable advice and comments. Thanks are also due to Dr. M. Moore for critical reading of the manuscript. This work was partly supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan; by a grant from National Center for Nervous, Mental and Muscular Disorders of the Ministry of Health and Welfare, Japan; and by funds obtained under the Life Science Project from the Institute of Physical and Chemical Research, Japan.

1. Takemori, S. & Kominami, S. (1984) *Trends Biol. Sci.* **9**, 393–396.
2. New, M. I. & Levine, L. S. (1984) in *Pediatric Adolescent Endocrinology*, eds. New, M. I. & Levine, L. S. (Karger, Basel, Switzerland), Vol. 13, pp. 1–46.
3. Carroll, M. C., Campbell, R. D. & Porter, R. R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 521–525.
4. White, P. C., New, M. I. & Dupont, B. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7505–7509.
5. White, P. C., Grossberger, D., Onufer, B. J., Chaplin, D. D., New, M. I., Dupont, B. & Strominger, J. L. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1089–1093.
6. Belt, K. T., Carroll, M. C. & Porter, R. R. (1984) *Cell* **36**, 907–914.
7. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
8. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
9. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
10. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
11. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–321.
12. Reynolds, G. A., Basu, S. K., Osborne, T. F., Chin, D. J., Gil, G., Brown, M. S., Goldstein, J. L. & Lusky, K. L. (1985) *Cell* **38**, 275–285.
13. White, P. C., Chaplin, D. D., Weis, J. H., Dupont, B., New, M. I. & Seidman, J. G. (1984) *Nature (London)* **312**, 465–467.
14. Chung, B.-C., Matteson, K. J. & Miller, W. L. (1985) *DNA* **4**, 211–219.
15. Hood, L., Steinmetz, M. & Malissen, B. (1983) *Annu. Rev. Immunol.* **1**, 529–568.
16. Slightom, J. L., Blechl, A. E. & Smithies, O. (1980) *Cell* **21**, 627–638.
17. Nonaka, M., Takahashi, M., Natsuume-Sakai, S., Nonaka, M., Tanaka, S., Shimizu, A. & Honjo, T. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6822–6826.
18. Suwa, Y., Mizukami, Y., Sogawa, K. & Fujii-Kuriyama, Y. (1985) *J. Biol. Chem.* **260**, 7980–7984.
19. Mizukami, Y., Sogawa, K., Suwa, Y., Muramatsu, M. & Fujii-Kuriyama, Y. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3958–3962.
20. Sogawa, K., Gotoh, O., Kawajiri, K., Harada, T. & Fujii-Kuriyama, Y. (1985) *J. Biol. Chem.* **260**, 5026–5032.
21. Sogawa, K., Gotoh, O., Kawajiri, K. & Fujii-Kuriyama, Y. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5066–5071.
22. Hines, R. N., Levy, J. B., Shen, M.-L., Renli, A. M. & Bresnik, E. (1985) *Arch. Biochem. Biophys.* **237**, 465–476.
23. Gonzalez, F. J., Kimura, S. & Nebert, D. W. (1985) *J. Biol. Chem.* **260**, 5040–5049.
24. Kioussis, D., Eiferman, F., ven de Rijn, P., Gorin, M. B., Ingram, R. S. & Tilghman, S. M. (1981) *J. Biol. Chem.* **256**, 1960–1967.
25. Sharp, P. A. (1985) *Cell* **42**, 397–400.