

A Novel Polymorphic Cytochrome P450 Formed by Splicing of CYP3A7 and the Pseudogene CYP3AP1*

Received for publication, March 1, 2005, and in revised form, May 9, 2005
Published, JBC Papers in Press, June 2, 2005, DOI 10.1074/jbc.M502309200

Cristina Rodriguez-Antona^{‡§}, Magnus Axelson[¶], Charlotta Otter^{||}, Anders Rane^{**},
and Magnus Ingelman-Sundberg[‡]

From the [‡]Division of Molecular Toxicology, Institute of Environmental Medicine, Karolinska Institute, 17177 Stockholm, the [¶]Department of Clinical Chemistry, Karolinska University Hospital, 171776 Stockholm, ^{||}Molecular Biology, AstraZeneca Research and Development, 43183 Mölndal, Sweden, and the ^{**}Karolinska Institute at Huddinge University Hospital, 14186 Stockholm, Sweden

The cytochrome P450 3A7 (CYP3A7) is the most abundant CYP in human liver during fetal development and first months of postnatal age, playing an important role in the metabolism of endogenous hormones, drugs, differentiation factors, and potentially toxic and teratogenic substrates. Here we describe and characterize a novel enzyme, CYP3A7.1L, encompassing the CYP3A7.1 protein with the last four carboxyl-terminal amino acids replaced by a unique sequence of 36 amino acids, generated by splicing of CYP3A7 with CYP3AP1 RNA. The corresponding CYP3A7-3AP1 mRNA had a significant expression in liver, kidney, and gastrointestinal tract, and its presence was found to be tissue-specific and dependent on the developmental stage. Heterologous expression in yeast revealed that CYP3A7.1L was a functional enzyme with a specific activity similar to that of CYP3A7.1 and, in some conditions, a different hydroxylation specificity than CYP3A7.1 using dehydroepiandrosterone as a substrate. CYP3A7.1L was found to be polymorphic due to a mutation at position –6 of the first splicing site of CYP3AP1 (CYP3A7_39256T→A), which abrogates the pseudogene splicing. This polymorphism had pronounced interethnic differences and was in linkage disequilibrium with other functional polymorphisms described in the CYP3A locus: CYP3A7*2 and CYP3A5*1. Therefore, the resulting CYP3A haplotypes express different sets of enzymes within the population. In conclusion, a novel mechanism, consisting of the splicing of the pseudogene CYP3AP1 to CYP3A7, causes the formation of the novel CYP3A7.1L having a different tissue distribution and functional properties than the parent CYP3A7 enzyme, with possible developmental, physiological, and toxicological consequences.

The cytochrome P450s (CYPs)¹ of families 1–3 are the most prominent enzymes catalyzing the biotransformation of exoge-

nous compounds in the liver and have a major role in drug metabolism. The CYP3A subfamily has a special relevance because CYP3A7 and CYP3A4 metabolize a large number of therapeutic drugs and are the most abundant CYPs in fetal and adult liver, respectively. The human CYP3A locus is located on chromosome 7q21-q22.1 and contains four genes, CYP3A4, CYP3A5, CYP3A7, and CYP3A43, and three pseudogenes, CYP3AP1, CYP3AP2, and CYP3AP3 (1, 2). The CYP3A genes were acquired by gene duplications and divergence of the proteins during evolution, generating enzymes with different functions and expression patterns. The pseudogenes probably represent recombination events followed by deletions resulting in a non-functional combination of exons: exons 1, 2, and 13 (CYP3AP1), exons 1 and 2 (CYP3AP2), and exon 1 (CYP3AP3) (1).

Developmental- and tissue-specific mechanisms regulate the expression of the CYP3A enzymes. Among adults, CYP3A4 is the dominant CYP in liver and intestines, being involved in the metabolism of more than 50% of currently used therapeutic drugs (3). CYP3A5 can be present at relevant levels in both fetal and adult liver, although in a polymorphic manner (4). The low expression of CYP3A43 and improper folding in mammalian systems rule out any significant contribution of this enzyme to CYP3A activity (5, 6). In fetal liver, where the number of expressed xenobiotic metabolizing enzymes is small, CYP3A7 is present at high levels (7–9). CYP3A7 plays an important role in the metabolism of key steroids in the adrenals and gonads, of potentially toxic and teratogenic endogenous substrates such as retinoic acid (10–13), and of many xenobiotics that reach the fetus, such as therapeutic drugs for women and substances of abuse (14–17).

The important role of CYP3A7 in the metabolism of endogenous and exogenous compounds, added to the fact that most drug-metabolizing P450 genes are not expressed during fetal stages, results in a well conserved gene with few allelic variants (18). In contrast, the adult liver contains multiple CYPs with different substrate specificities to facilitate the metabolism and elimination of a wide range of compounds. Here we describe a mechanism that generates a novel CYP3A7 enzyme (CYP3A7.1L) by splicing between the pseudogene CYP3AP1 and CYP3A7. The novel mRNA, containing 13 exons of CYP3A7 and 2 additional exons from CYP3AP1 at the 3' end (CYP3A7-3AP1, GenBankTM accession number AF315325) (1) (see Fig. 1), is found at relevant amounts in several tissues and encodes a functional enzyme.

EXPERIMENTAL PROCEDURES

Materials—Protease inhibitor mixture tablets were from Roche Applied Science. Dehydroepiandrosterone (DHEA), 16 α -hydroxy-DHEA, glucose, galactose, and NADPH were purchased from Sigma. [1,2,6,7-

* This work was supported by a grant from the Swedish Research Council, National Institutes of Health (Grant NIGMS 1-R01 GM60548) and by a Fellowship of the European Community programme Quality of Life and Management of Living Resources under Contract Number QL65-CT-2002-51733 (to C. R.-A). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: Division of Molecular Toxicology, Institute of Environmental Medicine, Karolinska Institutet, Nobels väg 13, 17177 Stockholm, Sweden. Tel.: 46-8-52487711; Fax: 46-8-337327; E-mail: Cristina.Rodriguez-Antona@imm.ki.se.

¹ The abbreviations used are: CYP, cytochrome P450; DHEA, dehydroepiandrosterone; GC/MS, gas chromatography mass spectrometry; HPLC, high performance liquid chromatography.

³H(N)]DHEA (specific activity 74Ci/mmol) was purchased from PerkinElmer Life Sciences.

Liver Tissue—A panel of 9 fetal and 15 adult human liver pieces frozen in liquid nitrogen and stored at -70°C was used to isolate total RNA and genomic DNA and to prepare microsomes by subcellular fractionation. Fetal human liver pieces were obtained from legal abortions on sociomedical indications. The study was approved by the Ethics Committee at Karolinska Institutet, Stockholm, Sweden. The characteristics of the adult livers and the data concerning gestational ages of the fetal livers have been described previously (18, 19). Human liver microsomes were prepared by subcellular fractionation as described elsewhere (20).

Real-time Quantitative PCR—Total RNA was extracted from the 15 human adult livers and from the 9 fetal human livers using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Two μg of total RNA were reverse-transcribed using SuperScript II reverse transcriptase (Invitrogen) and an oligo(dT14) primer. Quantitative real-time PCR assay was carried out with the ABI Prism 7700 (PE Applied Biosystems, Foster City, CA) and using the SyberGreen PCR master mix (PE Applied Biosystems). The following primers were used: CYP3A7, forward 5'-CTATGATACGTGCTACAGT-3' and reverse 5'-TCAGGCTCCACTTACGGTCT-3' primers; CYP3A7-3AP1, forward 5'-TGATAGATATGGAGCCATT-3' and reverse 5'-GTCTCTGGTGTCT-AGGT-3' primers; β -actin, forward 5'-CGTACCCTGGCAGTCTGAT-3' and reverse 5'-GTGTTGGCGTACAGGCTTTGCG-3' primers. Aliquots of the PCR reactions were subjected to electrophoresis for size and purity confirmation. Standard curves were constructed with serial 10-fold dilutions of an accurately determined concentration of a DNA fragment containing the cDNA of interest. Normalization for variation in reverse transcriptase efficiency and small differences in the amount of RNA added to each reaction was carried out with the internal standard β -actin.

Sequencing of CYP3AP1 and Genotyping—The region of CYP3AP1 with high similarity to CYP3A7 intron 1-exon 2 was analyzed by direct sequencing using forward 5'-TCAGTGACTTACCAGCCCC-3' and reverse 5'-TTCAACTGAGGCAAACCTGA-3' primers and the ABI Prism BigDye terminator cycle sequencing ready reaction kit and analyzed on an ABI Prism 377 DNA sequencer (Applied Biosystems). A T \rightarrow A change in CYP3AP1 (CYP3A7_39256 T \rightarrow A, the number indicating the position of the polymorphism with respect to the CYP3A7 translational start site in the GenBankTM reference sequence AF280107, was detected in the two fetal livers that did not have CYP3A7-3AP1 mRNA. A method for genotyping CYP3A7_39256 A was set up using the primer pair: forward 5'-TCAGTGACTTACCAGCCCC-3' and reverse 5'-TTCAACTGAGGCAAACCTGA-3' in a PCR reaction starting with 1 min at 95°C followed by 36 cycles of 20 s at 95°C , 30 s at 57°C , 40 s at 72°C , and a final extension of 5 min at 72°C and amplification of a 369-bp DNA fragment. This PCR product was diluted 100 times and served as a template for a subsequent PCR reaction that introduced a RcaI restriction site in the A-allele (CYP3A7_39256 A) using 5'-TCATAAAGTCACAATCCTGAGACCTGATTCATG-3' forward primer, with the two mismatches introduced underlined, and 5'-GCCAAA-GAGTGAGCTCAAAA-3' reverse primer. These primers amplified a 170-bp DNA by a PCR of 1 min at 95°C followed by 27 cycles of 15 s at 95°C , 20 s at 55°C , 40 s at 72°C , and a final extension of 5 min at 72°C . Ten μl of this PCR fragment were incubated with RcaI overnight at 37°C and separated by electrophoresis in a 3% high resolution agarose gel. The CYP3A7_39256 T allele resulted in the 170-bp undigested PCR fragment, whereas CYP3A7_39256 A resulted in two DNA fragments of 136 and 34 bp. Genotyping of CYP3A7*2 was performed as described previously (18).

Subjects—Allelic frequencies were determined using genomic DNA from unrelated individuals from different populations. Sixty-five volunteers from the Zaragoza area in the north of Spain (21), 45 Swedish (19, 22) individuals, and 96 Chinese individuals (23) were included. The genomic DNA was isolated from blood except for the Swedish population where it was extracted from liver samples using the QIAamp tissue kit (Qiagen). The present study was approved by the Ethics Committee at Karolinska Institutet.

CYP3A7-3AP1 Expression Pattern Analysis—cDNAs from different tissues were obtained from BD Biosciences (human MTC panel I and II, human fetal MTC panel, and human digestive system MTC panel) and were subjected to PCR amplification using the primers CYP3A7, 5'-TTCCAAGCTATGTTCTTCAT-3' and 5'-CATTTCAGGGTCTATTTGT-3' and CYP3A7-3AP1, 5'-TGATAGATATGGAGCCATT-3' and 5'-GTCTCTGGTGTCTAGGT-3' with 1 min at 95°C followed by 40 cycles of 20 s at 95°C , 30 s at 57°C , 40 s at 72°C , and a final extension of 7 min at 72°C .

Heterologous Expression of CYP3A7.1 and CYP3A7.1L in Yeast—Full-length CYP3A7 and CYP3A7-3AP1 coding regions were PCR-amplified from human liver cDNA using Elongase (Invitrogen). The forward primer was the same for CYP3A7 and CYP3A7-3AP1: 5'-AAGGGGATC-CAAAATGGATCTCATCCCAAACCTGGCCG-3' and introduced three A residues in front of the translation start to increase the expression efficiency in yeast (24) and a BamHI site at the 5' end. The reverse primers contained a KpnI site at the 5' end: CYP3A7, 5'-CTTAGGTAC-CTCAGGCTCCACTTACGGTCT-3' and CYP3A7-3AP1, 5'-AAGCGG-TACCTCAGCGGGATCTGATGGTAG-3'. The underlined nucleotides correspond to the restriction enzyme sites. After double digestion of the amplified DNA with BamHI/KpnI, the fragments were cloned into the pYeDP60 expression vector (25). Plasmids from positive transformed bacterial colonies were isolated, and the sequence of the insert was examined using the ABI Prism dye terminator cycle sequencing ready reaction kit (Applied Biosystems) and analyzed on an ABI Prism 377 DNA sequencer (Applied Biosystems).

The *Saccharomyces cerevisiae* strain W(R) was constructed by substitution of the natural promoter of the cytochrome P450 reductase (CPR1) gene by the galactose-inducible and glucose-repressed promoter GAL10-CYC1 in the strain W303-1B: MATa ade2-1 his3-11-15 leu2-3, 112 trp1-1 ura3-1 (26, 27), and the strain INVSc1-HR: MAT a his3A1 leu2 trp1-289 ura3-52 (pFL-35 human reductase) expressing human NADPH-cytochrome P450 reductase was a gift from the LINK project (a program of the University of Dundee/Biotechnology and Biology Research Council/Department of Trade and Industry/Pharmaceutical Industry). Both strains W(R) and INVSc1-HR were used to express CYP3A7.1 and CYP3A7.1L, essentially as described elsewhere (26, 28). In brief, the yeast cells were transformed with the pYeDP60 expression vectors containing a galactose-inducible promoter in front of the cDNA encoding CYP3A7.1 or CYP3A7.1L. To achieve high expression levels, the yeast cells were grown to high density with glucose as the main energy source; thereafter, galactose was added to induce expression. After expression, the cells were harvested and mechanically disrupted as described previously (28). Microsomes containing the recombinant enzyme were thereafter prepared by differential centrifugation ($20,000 \times g$ for 10 min plus $100,000 \times g$ for 60 min), and protein concentration was determined according to Bradford (29). Production of active CYP3A7.1L in W(R) cells proved to be difficult, and different growth and induction conditions were tested. A successful CYP3A7.1L overexpression was achieved at 28°C and 7 h of galactose induction. The W(R) data in Table II was generated using this preparation of CYP3A7.1L and the parallel preparations for CYP3A7.1 and control yeast transformed with empty vector.

Immunoblot Analysis—Microsomal proteins were electrophoresed in SDS-polyacrylamide gels using the Mini-PROTEAN II electrophoresis cell (Bio-Rad), transferred to Hybond-C extra nitrocellulose membranes (Amersham Biosciences), and incubated with a CYP3A antibody (Gen-test, catalog numbers A254 and 458223) following the manufacturer's instructions. Both antibodies, A254 and 458223, recognize CYP3A4, CYP3A5, CYP3A7, and CYP3A7.1L (this latter, according to our data using heterologous expression systems). A254 is a monoclonal antibody, and 458223 is polyclonal antibody produced by using rat CYP3A2 as immunogen. After washing, the membranes were incubated with the corresponding horseradish peroxidase-conjugated secondary antibody, and the proteins were visualized using the SuperSignal West Pico chemiluminescence method (Pierce) and scanned using LAS-1000 (FujiFilm, Dusseldorf, Germany). The relative intensity of each band as determined by Image Gauge V3.6. The CYP3A7.1 and CYP3A7.1L content was determined from standard curves derived from a microsomal sample quantified with respect to the apo-CYP3A7 content determined by peptide-specific antibody and peptide-conjugated lysozyme, kindly provided by Robert J. Edwards, Imperial College London, UK, and Sarah C. Sim, Karolinska Institutet.²

Dehydroepiandrosterone Hydroxylation Assay and Identification of Metabolites—Measurements of DHEA activity were performed in a final volume of 0.5 ml with 100 mM potassium phosphate buffer, pH 7.4, containing 15–70 nM [³H]DHEA and between 250 and 500 μg of yeast microsomal protein. The reaction mixtures were preincubated at 37°C for 5 min before the addition of NADPH to a final concentration of 2 mM. The reactions were carried out at 37°C for 30–60 min and stopped by the addition of 5 ml of dichloromethane, and the extraction of the metabolites was carried out by vortexing for 20 min, 10 min of centrifugation at $5000 \times g$, and separation and drying of the organic phase

² S. C. Sim, R. J. Edwards, A. R. Boobis, and M. Ingelman-Sundberg (2005) *Pharmacogenetics*, in press.

under a stream of N_2 gas. Dried extracts were dissolved in 100 μ l of mobile phase and analyzed by high-performance liquid chromatography (HPLC). A 5 μ m 4×250 mm LiChrospher-100 RP-18 column (Merck) using a 64:36 methanol:H₂O (v/v) isocratic mobile phase was used to separate the hydroxylated products. Radioactive peaks were detected by an in-line Berthold radioactivity monitor LB506C-1, and a Berthold LB5035 pump was used for the scintillation liquid at a 1.5 ml/min flow rate (Quicksafe Flow 2, Zinsser Analytic). Values greater than 0.1 pmol of metabolite were accurately detected with this method. Linearity with protein and time was observed for the production of DHEA metabolites. It should be noted that the 3H label in position C-7 of DHEA will result a somewhat diminished signal of the 7-hydroxylated metabolites.

For identification of the DHEA metabolites by gas chromatography mass spectrometry (GC/MS), corresponding incubation mixtures contained 50–150 μ M unlabeled DHEA, and the reactions were stopped by the addition of FOLCH solution (CHCl₃:MeOH, 2:1). After vortexing, the organic phase was collected and dried under a gentle stream of N_2 . The residue was then dissolved in 3 ml of methanol followed by the addition of 2 ml of water. The solution was passed through a Sep-Pak C₁₈ cartridge containing octadecylsilane-bonded silica (Waters Associates Inc., Milford, MA) followed by 5 ml of water. The total effluent was collected, and the organic solvent was removed *in vacuo*. The remaining aqueous phase was then passed through the same unwashed Sep-Pak C₁₈ cartridge again, prior to a wash with 5 ml of water. Steroids were then eluted with 10 ml of 75% aqueous methanol. This solution was passed through a column (40 \times 4 mm) of the strong lipophilic anion exchanger, triethylaminoxypropyl Sephadex LH-20, in bicarbonate form (30). After elution with an additional 5 ml of methanol, the total effluent from the column was taken to dryness *in vacuo*. The residue was then transferred with methanol to a stoppered tube and dried under N_2 , and steroids were trimethylsilylated in 0.2 ml of pyridine/hexamethyldisilazane/trimethylchlorosilane, 3:2:1 (by volume), by heating at 60 $^{\circ}C$ for 30 min. The reagents were removed under N_2 , and the derivatives were redissolved in 50 μ l of hexane. GC/MS was performed using a Finnigan SSQ 710 instrument housing a fused silica column (25 m \times 0.32 mm) coated with a 0.17- μ m layer of cross-linked methyl silicone (Ultra 1, Hewlett-Packard, Palo Alto, CA) ending in the ion source. An on-column injection device was used. The oven temperature was 50 $^{\circ}C$ during the injection and, after 3 min, it was rapidly increased to 185 $^{\circ}C$ and then programmed to 280 $^{\circ}C$ at a rate of 5 $^{\circ}C$ /min. The electron energy was 50 eV, and repetitive scanning (30 scans/min) over the *m/z* range of 50–800 was started after a suitable delay. The identification of a steroid was based on the retention time and complete mass spectrum, which were compared with those of the authentic steroid. The retention indices (Kovats) for DHEA, 7 α -hydroxy-DHEA, 7 β -hydroxy-DHEA, and 16 α -hydroxy-DHEA were 2515, 2590, 2670, and 2725, respectively, and mass spectra of the two 7-hydroxylated DHEA metabolites showed one intense ion (*m/z* 358, molecular ion minus 90), whereas the major significant ions in the spectrum of 16 α -hydroxy-DHEA were *m/z* 448 (molecular ion), 304, 214, and 129.

Statistical Analysis—Data were analyzed using GraphPad InStat version 3.00 for Windows 95, GraphPad Software, San Diego, CA. Statistical analysis was done by the Mann-Whitney test. Results are expressed as mean \pm S.D. Differences were considered significant when *p* values were less than 0.05.

RESULTS

Expression of CYP3A7-3AP1 mRNA in Human Liver—To search for polymorphic variants, a forward primer corresponding to CYP3A4/3A7 exon 12 and a reverse primer in the 3'-untranslated region CYP3A4 were designed to amplify CYP3A4 cDNA by reverse transcription-PCR in a panel of 15 adult and 9 fetal human livers. However, a DNA fragment of a different molecular weight was unexpectedly amplified. Sequencing of this DNA revealed that it was a CYP3A7 splicing variant that consisted of two fragments of the pseudogene CYP3AP1, which are highly similar to exons 2 and 13 of CYP3A7, joined to CYP3A7 RNA at exon 13 in an alternative splice site, similarly to the one described by Finta and Zaphiropoulos (1) (Fig. 1A). All exons of this novel transcript are flanked by canonical splicing sites, which strongly suggests that it is the result of an RNA splicing event. CYP3A7, similarly to the other CYP3A genes, contains 13 exons, whereas the novel CYP3A7-3AP1 mRNA contains 15 exons, resulting in a longer CYP3A7 protein (CYP3A7.1L). The translational stop

codon is located seven nucleotides downstream in the last exon from the pseudogene. In addition, the reading frame of the pseudogene regions is shifted, and the last four amino acids of CYP3A7 are replaced by a unique 36-amino-acid sequence (Fig. 1B).

As shown in Fig. 2, the expression levels of CYP3A7-3AP1 mRNA, as determined by a quantitative reverse transcription-PCR technique, was higher in fetal than in adult livers, with the exception of two fetal samples (FL61 and FL65) having extremely low CYP3A7-3AP1 mRNA levels. Using two CYP3A-specific antibodies, CYP3A7.1L was detected by Western blotting in liver samples with CYP3A7-3AP1 mRNA expression (Fig. 2A, inset; FL65, which lacks CYP3A7.1L, is used as a negative control).

To explore the molecular mechanisms responsible for the alternative splicing, CYP3A7 mRNA levels were measured and compared with the expression of CYP3A7-3AP1 mRNA. The correlation between CYP3A7-3AP1 and CYP3A7 mRNA of 22 livers, all with exception of FL61 and FL65, was high ($r_2 = 0.94$; $r = 0.97$; $p < 0.00001$). On average, the CYP3A7-3AP1 mRNA levels were 1% of those of CYP3A7 mRNA, but substantial differences were noted between fetal and adult livers, with mean ratios of 0.5 and 1.2%, respectively, indicating that the splicing of the pseudogene was favored in the adult livers. Interestingly, the two fetal livers with extremely low CYP3A7-3AP1 mRNA expression (FL61 and FL65) had intermediate to high expression of CYP3A7 mRNA, suggesting a genetic defect preventing the splicing of CYP3AP1.

CYP3A7_39256 T→A Prevents CYP3AP1 Splicing—Sequencing of genomic DNA from the fetal livers lacking CYP3A7-3AP1 mRNA revealed a T→A change at -6 of the first acceptor splicing site of CYP3AP1. The splicing acceptor site of CYP3A7 intron 1 and the similar sequence of CYP3AP1, together with the nucleotide change, are shown in Fig. 3. Interestingly, the C at -8 of CYP3A7 intron 1 is a G in CYP3AP1. Thus, this nucleotide change at the pyrimidine stretch seems to be sufficient to affect the splicing and to shift it to an alternative site in CYP3AP1. This suggested that the changes in the pyrimidine sequence of intron 1 have a significant impact on the splicing efficiency. Thus, adult livers homozygous for T at position 39256 had a higher CYP3A7-3AP1/CYP3A7 ratio than those heterozygous in this position ($p = 0.003$; Fig. 2C), although the amounts of CYP3A7-3AP1 were still low when compared with CYP3A7 mRNA. Similarly, fetal livers homozygous for A did not express CYP3A7-3AP1 mRNA. Thus, the data indicated that the CYP3A7_39256 T→A change at position -6 of the pseudogene prevents the splicing. When the CYP3A7-3AP1/CYP3A7 ratio of adult and fetal livers homozygous for CYP3A7_39256 T/T were compared, there was also a significant difference in the expression (1.4% versus 0.5%, respectively, $p = 0.001$), further indicating that the splicing of the pseudogene is regulated in a developmental-specific manner.

Similarly to CYP3A7 and CYP3AP1, a pseudogene containing an intron 1/exon 2 region with conserved splice sites is located downstream of CYP3A4 (CYP3AP2). However, no alternatively spliced RNA could be amplified by reverse transcription-PCR using primers at CYP3A4 exon 13 and at CYP3AP2 exon 2 with human fetal and adult liver RNA reverse-transcribed with poly(dT) or a pseudogene-exon 2-specific primer (data not shown). The cause might be inherent in the lack of an exon 13-like sequence in CYP3AP2 and, therefore, of a polyadenylation site.

CYP3A7_39256 A Frequency and Linkage Disequilibrium with CYP3A7*2 and CYP3A5*1—The frequency of CYP3A7_39256 A showed large interethnic differences: 8, 28, and 59% in Caucasians, Chinese, and African Americans, re-

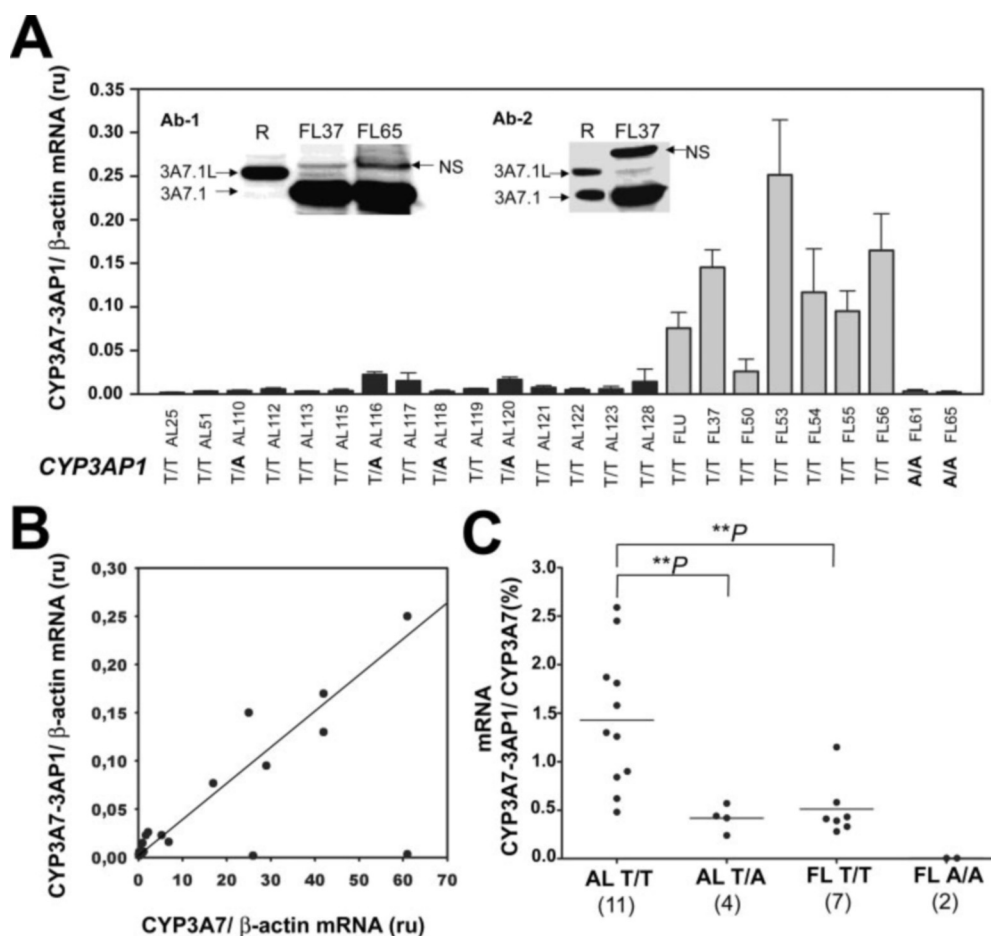


FIG. 2. CYP3A7-3AP1 mRNA and CYP3A7.1L expression in human liver. *A*, the content of CYP3A7-3AP1 mRNA was measured in 15 adult (AL) and 9 fetal (FL) human livers by quantitative real-time PCR, and the amounts were normalized by the content of β -actin. The CYP3A7 genotype (CYP3A7_39256 A or CYP3A7_39256 T) of the different samples is shown. *Inset*, CYP3A7.1L is immunodetected in a fetal liver microsomal sample with CYP3A7-3AP1 mRNA expression (FL37) but not in FL65, which lacks CYP3A7-3AP1 expression. The CYP3A-specific antibodies used for Western blot analysis are described under "Experimental Procedures." Ab-1 corresponds to A254, and Ab-2 is 458223. In fetal livers, the mean ratio CYP3A7-3AP1/CYP3A7 mRNA is 0.5%, if an identical ratio is assumed at protein level (identical translation efficiency and stability of CYP3A7.1 and CYP3A7.1L proteins); 2 pmols of CYP3A7.1L/mg of microsomal protein is reasonable in fetal liver (18). R, recombinant; NS, nonspecific. *B*, correlation between CYP3A7-3AP1 and CYP3A7 mRNA. When all livers, except FL61 and FL65 with no quantifiable CYP3A7-3AP1 mRNA, were analyzed, the correlation was: $r^2 = 0.94$; $r = 0.97$; $p < 0.00001$. ru, relative units. *C*, differences in CYP3A7-3AP1/CYP3A7 mRNA ratio between adult livers and fetal livers with different CYP3A7 genotypes. The mean value in the different groups is indicated with a horizontal line, and the number of livers in each group is shown between brackets. **, $p < 0.01$, significantly different expression with respect to adult CYP3A7_39256 T/T livers (AL T/T) using Mann-Whitney test.

intron 1- EXON 2

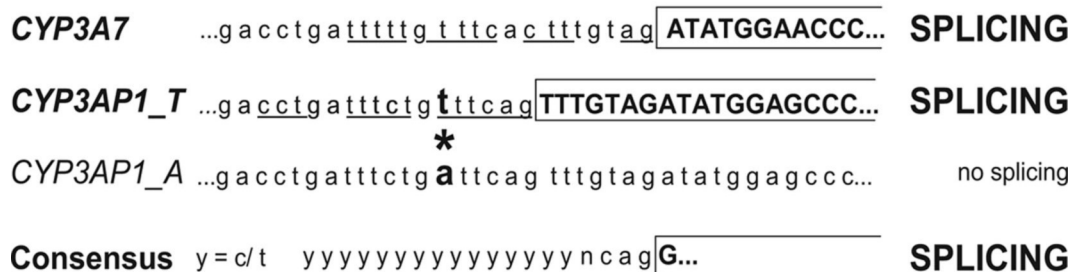


FIG. 3. Intron 1-exon 2 boundaries in CYP3A7 and CYP3A7. Intrinsic sequences are in lowercase, and exons are shown in uppercase. The consensus splicing acceptor site is shown at the bottom. The nucleotides that are identical to the consensus splicing sequence are underlined. The polymorphism in CYP3A7 that abolishes the pseudogene splicing (CYP3A7_39256 T→A) is marked with a star.

onstrating that CYP3A7.1L is a functional enzyme. In W(R) cells 7 α -, 16 α -, and 7 β -OH DHEA were produced (Fig. 5), and in INVSc1HR cells, only 16 α -OH DHEA could be detected (Table II). The difference in the production of metabolites could be caused by different protein-protein interactions of CYP3A7.1L

with the overexpressed P450 reductases. Western blotting analysis showed that the amounts of CYP3A7.1L and CYP3A7.1 in W(R) and INVSc1HR strains were different, with a 50-fold lower CYP3A7.1L/CYP3A7.1 ratio in the INVSc1HR cells (Table II). This is likely a result of different interactions of

TABLE I
Allele frequencies of CYP3A7_39256 A and CYP3A7*2 in different ethnic groups and linkage disequilibrium between them

Ethnic group	No. of chromosomes analyzed	Allele frequency ^a		Haplotype frequency ^b	
		CYP3A7_39256 A	CYP3A7*2	CYP3A7*1A	CYP3A7*2A
Caucasians					
Spanish	130	0.07	0.07	0.93	0.07
Swedish	90	0.09	0.09	0.91	0.09
Chinese	192	0.28	0.28	0.72	0.28
Africans	46 ^c /180 ^d	0.59 ^c	0.62 ^d		

^a The frequency of the allele was calculated by using the formula: frequency = [2*(number of people homozygous for the allele) + (number of heterozygous people)] / [2*(total number of people)]. CYP3A7*1 and CYP3A7_39256 T, GenBank™ accession number AF280107.

^b The frequencies of the haplotypes are calculated for each population. Haplotypes and allele combinations are: CYP3A7*1A (CYP3A7*1/CYP3A7_39256 T; proteins: CYP3A7.1 and CYP3A7.1L), CYP3A7*2A (CYP3A7*2/CYP3A7_39256 A; proteins: CYP3A7.2).

^c Data from NCBI, database of single nucleotide polymorphisms (dbSNP) build 124. CYP3A7_39256 T>A or rs2740565 was genotyped in 23 samples of African American descent from the Coriell Cell Repository. For European American and Chinese descent a 0.04 and 0.15 allele frequency is found using 24 samples from the Coriell Cell Repository.

^d Data from Rodriguez-Antona *et al.* (18).

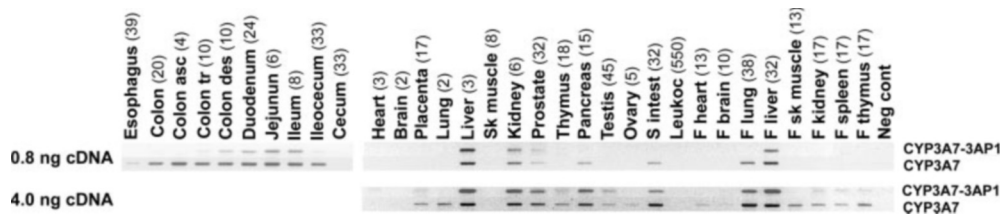


FIG. 4. Tissue distribution of CYP3A7-3AP1 mRNA. Adult and fetal human cDNA panels (BD Biosciences) were used to amplify by PCR the cDNAs of CYP3A7 and CYP3A7-3AP1, using 40 cycles of PCR, 0.8 or 4.0 ng of cDNA, and specific primers, as described under "Experimental Procedures." When 32 PCR cycles were used, the higher expression of CYP3A7 when compared with CYP3A7-3AP1 could be appreciated in adult and fetal liver (data not shown). The number in brackets indicates the number of pooled tissues, in all cases corresponding to Caucasians. The small intestine sample lacks mucosal lining. asc, ascendent; tr, transversal; des, descendent; Sk, skeletal; S intest, small intestine; Leukoc, leukocytes; F, fetal; Neg cont, negative control.

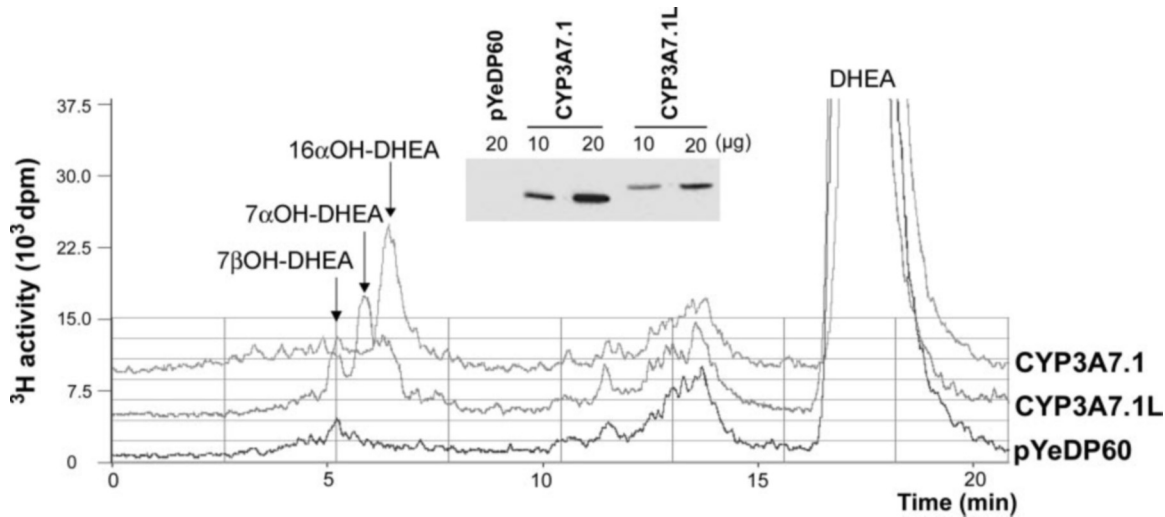


FIG. 5. HPLC metabolite profile of DHEA after incubation with microsomes from W(R) yeast expressing CYP3A7.1L and CYP3A7.1. The HPLC chromatograms resulting from DHEA incubations with microsomes from W(R) cells (overexpressing the yeast P450 reductase) overexpressing CYP3A7.1L or CYP3A7.1 and from W(R) cells transformed with empty vector (pYeDP60) are shown. Similar incubations were analyzed by GC/MS, and the different metabolites were identified. Inset, heterologous expression of CYP3A7.1 and CYP3A7.1L in W(R) cells. Ten and 20 μg of W(R) microsomal protein were analyzed by Western blot. The 32 extra amino acids of CYP3A7.1L carboxyl-terminal result in a protein of lower mobility than CYP3A7.1.

CYP3A7.1L with the yeast and human reductases or could reflect differences in CYP3A7.1L carboxyl-terminal folding and stability.

DISCUSSION

The highly polymorphic drug-metabolizing enzyme genes diversified through the coevolution of plants and animals. The differences in drug-metabolizing enzyme allele frequencies across human populations was suggested to reflect differences in xenobiotic environments mainly introduced by the adaptation to diet composition over thousands of years (31). Specifically, the wide range of structurally different substrates me-

tabolized by CYP3A enzymes, which is illustrated by their involvement in the biotransformation of more than 50% of all therapeutic drugs currently used, suggests that they might have been exposed to multiple selective pressures and that different variants may be the targets of these pressures. However, the key role of CYP3A4 and CYP3A7 in the metabolism of endogenous substrates, such as steroid hormones or bile acids (32, 33), imposed limitations to this variation and resulted in well conserved genes with rare deleterious mutations and low variability. In agreement with this, only two CYP3A7 enzymes have been described (CYP3A7.1 and CYP3A7.2) (18), both representing active enzymes. However, the pressure on these

TABLE II
Catalytic activity of CYP3A7.1 and CYP3A7.1L expressed in different yeast strains after incubation with DHEA

Yeast strain ^a	CYP3A7.1 expr. level	CYP3A7.1 activity (DHEA hydroxylation) ^b			CYP3A7.1L expr. level	CYP3A7.1L activity (DHEA hydroxylation)		
		7 β -OH	7 α -OH	16 α -OH		7 β -OH	7 α -OH	16 α -OH
	<i>pmol of P450^c</i>				<i>pmol of P450</i>			
W(R)-yeast reductase	39 \pm 13	— ^d	—	0.7 \pm 0.1	29 \pm 3	0.5 \pm 0.1	1.5 \pm 0.2	0.8 \pm 0.2
INVSc1-HR-human reductase	395 \pm 104	—	—	13.5 \pm 3.2	6 \pm 2	—	—	3.5 \pm 1.4

^a The *S. cerevisiae* strains used for heterologous expression are described under "Experimental Procedures." W(R) overexpress the yeast NADPH-cytochrome P450 reductase, and INVSc1-HR overexpress the human NADPH-cytochrome P450 reductase.

^b DHEA hydroxylation is expressed as pmol of metabolite/min/nmol P450. The amount of metabolites was quantified after subtraction of the background level found in control yeast transformed with empty vector.

^c Heterologous expression of CYP3A7.1 and CYP3A7.1L enzymes was measured by Western blot using reference samples as standard as described under "Experimental Procedures." The results are expressed as pmol of P450s per mg of microsomal protein.

^d —, below detection limit.

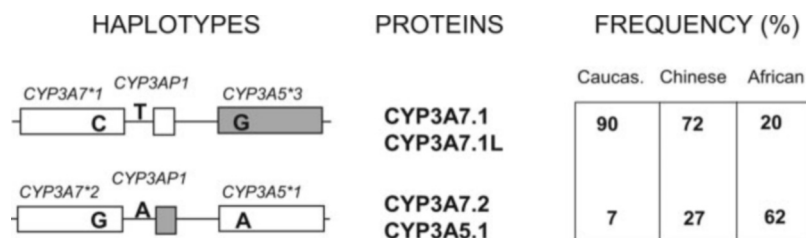


FIG. 6. Frequency of the different CYP3A haplotypes in the population. The two most frequent haplotypes that result in expression of different sets of CYP3A enzymes are depicted. In Caucasians and Chinese, the most common allele combination is *CYP3A7*1/CYP3A7_39256 T/CYP3A5*3*, which results in the expression of CYP3A7.1 and CYP3A7.1L proteins. In Africans, the most frequent allele combination, *CYP3A7*2/CYP3A7_39256 A/CYP3A5*1*, results in the expression of CYP3A7.2 and CYP3A5 proteins. The frequency of the different haplotypes is shown for the different populations. In Africans, the allele combination was calculated using the data in Rodriguez-Antona *et al.* (18) and assuming a total linkage disequilibrium between *CYP3A7*2* and *CYP3A7_39256 A*.

genes to diversify their functions is evidenced by mechanisms such as trans-splicing between CYP3A4, CYP3A5, and CYP3A43 mRNAs, which generates new enzymes without modification of the genes (34). In addition, this work showed that a novel CYP3A7-derived enzyme (CYP3A7.1L), generated by an alternative splicing of *CYP3A7* with *CYP3AP1*, is a functional enzyme. The corresponding mRNA, containing almost full sequence of *CYP3A7* plus two exons from regions of the *CYP3AP1* pseudogene (*CYP3A7-3AP1*), was found mainly in liver, kidney, and gastrointestinal tract. Interestingly, we found that the splicing efficiency between *CYP3A7* and *CYP3AP1* was regulated in a developmental- and tissue-specific manner (Fig. 2C, expression differences between fetal and adult liver and Fig. 4, tissue differences). This opened the possibility that CYP3A7.1L could be the major splicing form in certain tissues or at specific developmental stages.

A single nucleotide polymorphism at -6 of the first acceptor splice site of *CYP3AP1* (*CYP3A7_39256 T*→*A*) prevents the alternative splicing and results in no CYP3A7.1L expression (Fig. 3). Important interethnic differences in the frequency of *CYP3A7_39256 A* were found (Table I), being the predominant allele in Africans but less represented in Chinese and Caucasian populations. In addition, a linkage disequilibrium of *CYP3A7_39256 A* with *CYP3A7*2* and *CYP3A5*1* indicates that usually, specific sets of CYP3A enzymes are expressed with ethnic differences in their frequency. Two major haplotypes were identified: *CYP3A7*1/CYP3A7_39256 T/CYP3A5*3*, which results in the expression of CYP3A7.1 and CYP3A7.1L and is predominant in Caucasians and Chinese, and *CYP3A7*2/CYP3A7_39256 A/CYP3A5*1*, expressing CYP3A7.2 and CYP3A5 proteins and being the predominant in Africans (Fig. 6).

Expression of CYP3A7.1L in heterologous expression systems showed that CYP3A7.1L was a microsomal enzyme able to hydroxylate DHEA. Interestingly, the catalytic activity of the enzyme was shifted depending on the expression system used; in yeast overexpressing the yeast P450 reductase, DHEA was hydroxylated at 7 α -, 16 α -, and 7 β -positions, whereas in

yeast overexpressing human P450 reductase, DHEA was hydroxylated only at the 16 α -position. Miller *et al.* (35) showed that CYP3A4 catalyzes DHEA 7 β -, 16 α -, and 7 α -hydroxylations, whereas CYP3A7 hydroxylates DHEA at 16 α - and 7 β - but not at 7 α -positions. Similarly to Stevens *et al.* (8), they also showed that CYP3A7 was a more efficient 16 α -DHEA hydroxylase than CYP3A4 and that CYP3A7 was a less efficient 7 β -hydroxylase than CYP3A4. CYP3A7 and CYP3A4 share a 88% identity, and there is a large overlap in substrate specificity; however, there are differences in activity, probably caused by changes near the active site. The different hydroxylation specificity of CYP3A7.1L in the yeast systems, which does not occur for CYP3A7.1 (Table II), could be the result of different interactions of the longer carboxyl-terminal of CYP3A7.1L with the human and yeast reductases; alternatively, it could be caused by the different stability and catalytic activity of the enzyme.

Thompson *et al.* (36) showed that a relatively homogeneous class of haplotypes, spanning 150 kb, including *CYP3A4*, *CYP3A7* and *CYP3A5*, were driven to near fixation frequency in the non-African populations, suggesting the action of varying natural selective pressures. The authors proposed that lack of CYP3A5 protein (*CYP3A5*3*) conferred a selective advantage by influencing salt and water retention and risk for hypertension and could have been the target for the selection pressures in this locus. On the other hand, a different allele, in linkage disequilibrium with *CYP3A5*3*, could have conferred the selective advantage. The Caucasian most common haplotype consists, among other polymorphisms, of *CYP3A4*1A*, *CYP3A7*1*, *CYP3A7_39256 T*, and *CYP3A5*3*, whereas the African most common haplotype contains *CYP3A4*1B*, *CYP3A7*2*, *CYP3A7_39256 A*, and *CYP3A5*1*. An altered expression of *CYP3A4*1B* is controversial (19, 37), and CYP3A7.1 and CYP3A7.2 have differences in catalytic activity but are both functional enzymes (18). Therefore, it is unlikely that these alleles could have provided a significant adaptation advantage. On the other hand, *CYP3A7_39256 T* allele results in the

formation of a different CYP3A enzyme, which might have a different catalytic activity than CYP3A7 and has different expression patterns than CYP3A4. Therefore, *CYP3A7_39256 T* could have contributed to the complex evolutionary history of this genomic region. When chimpanzee (Ensembl, scaffold 37686) and human *CYP3A* locus are compared, a 99% identity between the coding regions of *CYP3A4*, *CYP3A5*, *CYP3A7*, and *CYP3A43* genes is found (38); however, the chimpanzee genome has five *CYP3A* genes. *CYP3A67*, with no human orthologue, is located between *CYP3A7* and *CYP3AP1*. The chimpanzee *CYP3AP1* has only two exons, 100% identical to the last human *CYP3AP1* exons, suggesting that a homologous recombination between *CYP3A67* and *CYP3AP1* resulted in loss of *CYP3A67* and the addition of exon 1 to the human *CYP3AP1*. The function and expression of *CYP3A67* is unknown, but the chimpanzee *CYP3AP1* corresponds to the T allele that permits alternative splicing, in this case with *CYP3A67*, and contains an extra nucleotide at the exon 2-like region that situates the two pseudogene exons in correct reading frame with *CYP3A67*.

Alternative splicing has recently emerged as a major source of protein diversity in mammalian genomes by allowing one single gene to encode functionally diverse proteins at precise time points and specific tissues. In this study, we showed that a novel enzyme is formed by the splicing of a pseudogene with the transcript of a closely located gene, a mechanism that is facilitated by the presence of consensus splicing sites in the pseudogene and about which we can speculate whether this splicing occurs in other loci. The expression of *CYP3A7.1L* is developmental- and tissue-specific, and in certain tissues, the *CYP3A7-3AP1* mRNA content is similar to that of *CYP3A7* mRNA. The highest expression of *CYP3A7.1L* is detected in fetal liver, where the presence of drug-metabolizing enzymes is very low when compared with the adult liver and where *CYP3A7* plays a key role in the metabolism of potentially toxic and teratogenic endogenous substrates such as retinoic acid and of xenobiotics reaching the fetus. These data, together with the fact that *CYP3A7* and *CYP3A7.1L* might have different catalytic activities, suggested that *CYP3A7.1L* could play an important role with developmental, physiological and toxicological processes.

Acknowledgments—We are indebted to Sarah C Sim and Robert Edwards for kindly providing us with a sample quantified for *CYP3A7*.

REFERENCES

- Finta, C., and Zaphiropoulos, P. G. (2000) *Gene (Amst.)* **260**, 13–23
- Gellner, K., Eiselt, R., Hustert, E., Arnold, H., Koch, I., Haberl, M., Deglmann, C. J., Burk, O., Buntfuss, D., Escher, S., Bishop, C., Koebe, H. G., Brinkmann, U., Klenk, H. P., Kleine, K., Meyer, U. A., and Wojnowski, L. (2001) *Pharmacogenetics* **11**, 111–121
- Bertz, R. J., and Granneman, G. R. (1997) *Clin. Pharmacokinet.* **32**, 210–258
- Kuehl, P., Zhang, J., Lin, Y., Lamba, J., Assem, M., Schuetz, J., Watkins, P. B., Daly, A., Wrighton, S. A., Hall, S. D., Maurel, P., Relling, M., Brimer, C., Yasuda, K., Venkataramanan, R., Strom, S., Thummel, K., Boguski, M. S., and Schuetz, E. (2001) *Nat. Genet.* **27**, 383–391
- Westlind, A., Malmbeo, S., Johansson, I., Otter, C., Andersson, T. B., Ingelman-Sundberg, M., and Oscarson, M. (2001) *Biochem. Biophys. Res. Commun.* **281**, 1349–1355
- Domanski, T. L., Finta, C., Halpert, J. R., and Zaphiropoulos, P. G. (2001) *Mol. Pharmacol.* **59**, 386–392
- Lacroix, D., Sonnier, M., Moncion, A., Cheron, G., and Cresteil, T. (1997) *Eur. J. Biochem.* **247**, 625–634
- Stevens, J. C., Hines, R. N., Gu, C., Koukouritaki, S. B., Manro, J. R., Tandler, P. J., and Zaya, M. J. (2003) *J. Pharmacol. Exp. Ther.* **307**, 573–582
- Burk, O., Tegude, H., Koch, I., Hustert, E., Wolbold, R., Glaeser, H., Klein, K., Fromm, M. F., Nuessler, A. K., Neuhaus, P., Zanger, U. M., Eichelbaum, M., and Wojnowski, L. (2002) *J. Biol. Chem.* **277**, 24280–24288
- Shimada, T., Yamazaki, H., Mimura, M., Wakamiya, N., Ueng, Y. F., Guengerich, F. P., and Inui, Y. (1996) *Drug Metab. Dispos.* **24**, 515–522
- Ohmori, S., Nakasa, H., Asanome, K., Kurose, Y., Ishii, I., Hosokawa, M., and Kitada, M. (1998) *Biochim. Biophys. Acta* **1380**, 297–304
- Lee, A. J., Conney, A. H., and Zhu, B. T. (2003) *Cancer Res.* **63**, 6532–6536
- Marill, J., Capron, C. C., Idres, N., and Chabot, G. G. (2002) *Biochem. Pharmacol.* **63**, 933–943
- Ladona, M. G., Gonzalez, M. L., Rane, A., Peter, R. M., and de la Torre, R. (2000) *Life. Sci.* **68**, 431–443
- Pellinen P, H. P., Stenback F, Niemitz M, Alhava E, Pelkonen, O., Lang, M. A., and Pasanen, M. (1994) *Eur. J. Pharmacol.* **270**, 35–43
- Matsunaga, T., Kishi, N., Higuchi, S., Watanabe, K., Ohshima, T., and Yamamoto, I. (2000) *Drug Metab. Dispos.* **28**, 1291–1296
- Foster, D. J., Somogyi, A. A., and Bochner, F. (1999) *Br. J. Clin. Pharmacol.* **47**, 403–412
- Rodriguez-Antona, C., Rane, A., and Ingelman-Sundberg, M. (2005) *Clin. Pharmacol. Ther.* **77**, 259–270
- Westlind, A., Lofberg, L., Tindberg, N., Andersson, T. B., and Ingelman-Sundberg, M. (1999) *Biochem. Biophys. Res. Commun.* **259**, 201–205
- Ernst, L., Siekevitz, P., and Palade, G. E. (1962) *J. Cell Biol.* **15**, 541
- Bernal, M. L., Sinues, B., Johansson, I., McLellan, R. A., Wennerholm, A., Dahl, M. L., Ingelman-Sundberg, M., and Bertilsson, L. (1999) *Pharmacogenetics* **9**, 657–660
- Westlind-Johansson, A., Malmbeo, S., Johansson, A., Otter, C., Andersson, T. B., Johansson, I., Edwards, R. J., Boobis, A. R., and Ingelman-Sundberg, M. (2003) *Drug Metab. Dispos.* **31**, 755–761
- Johansson, I., Yue, Q. Y., Dahl, M. L., Heim, M., Sawe, J., Bertilsson, L., Meyer, U. A., Sjoqvist, F., and Ingelman-Sundberg, M. (1991) *Eur. J. Clin. Pharmacol.* **40**, 553–556
- Krynetski, E. Y., Druitsa, V., Kovaleva, I. E., and Luzikov, V. N. (1995) *Pharmacogenetics* **5**, 103–109
- Urban, P., Cullin, C., and Pompon, D. (1990) *Biochimie (Paris)* **72**, 463–472
- Truan, G., Cullin, C., Reisdorf, P., Urban, P., and Pompon, D. (1993) *Gene (Amst.)* **125**, 49–55
- Bellamine, A., Gautier, J. C., Urban, P., and Pompon, D. (1994) *Eur. J. Biochem.* **225**, 1005–1013
- Bylund, J., Hidestrand, M., Ingelman-Sundberg, M., and Oliw, E. H. (2000) *J. Biol. Chem.* **275**, 21844–21849
- Bradford, M. M. (1976) *Cell* **72**, 248–254
- Axelsson, M., Ellis, E., Mörk, B., Garmark, K., Abrahamsson, A., Björkhem, I., Ericzon, B. G., and Einarsson, C. (2000) *Hepatology* **31**, 1305–1312
- Nebert, D. W. (1997) *Am. J. Hum. Genet.* **60**, 265–271
- Chang, T. K., Teixeira, J., Gil, G., and Waxman, D. J. (1993) *Biochem. J.* **291**, 429–433
- Xie, W., Radomska-Pandya, A., Shi, Y., Simon, C. M., Nelson, M. C., Ong, E. S., Waxman, D. J., and Evans, R. M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3375–3380
- Finta, C., and Zaphiropoulos, P. G. (2002) *J. Biol. Chem.* **277**, 5882–5890
- Miller, K. K., Cai, J., Ripp, S. L., Pierce, W. M. J., Rushmore, T. H., and Prough, R. A. (2004) *Drug Metab. Dispos.* **32**, 305–313
- Thompson, E. E., Kuttub-Boulos, H., Witonsky, D., Yang, L., Roe, B. A., and Di Rienzo, A. (2004) *Am. J. Hum. Genet.* **75**, 1059–1069
- Floyd, M. D., Gervasini, G., Masica, A. L., Mayo, G., George, A. L. J., Bhat, K., Kim, R. B., and Wilkinson, G. R. (2003) *Pharmacogenetics* **13**, 595–606
- Williams, E. T., Rodin, A. S., and Strobel, H. W. (2004) *Mol. Phylogenet. Evol.* **33**, 300–308