Identification and characterisation of novel polymorphisms in the *CYP2A* locus: implications for nicotine metabolism

Mikael Oscarson[a,*], Roman A. McLellan[a], Harriet Gullstén[b], José A.G. Agúndez[c], Julio Benítez[c], Arja Rautio[b], Hannu Raunio[b], Olavi Pelkonen[b], Magnus Ingelman-Sundberg[a]

[a] *Division of Molecular Toxicology, National Institute of Environmental Medicine, Karolinska Institutet, Box 210, 171 77 Stockholm, Sweden*
[b] *Department of Pharmacology and Toxicology, University of Oulu, Oulu, Finland*
[c] *Department of Pharmacology, Medical School, University of Extremadura, Badajoz, Spain*

Received 28 September 1999

Abstract The polymorphic human cytochrome P450 2A6 (CYP2A6) metabolises a number of drugs, activates a variety of precarcinogens and constitutes the major nicotine *C*-oxidase. A relationship between *CYP2A6* genotype and smoking habits, as well as incidence of lung cancer, has been proposed. Two defective alleles have hitherto been identified, one of which is very common in Asian populations. Among Caucasians, an additional defective and frequently distributed allele (*CYP2A6*3*) has been suggested to play a protective role against nicotine addiction and cigarette consumption. Here, we have re-evaluated the genotyping method used for the *CYP2A6*3* allele and found that a gene conversion in the 3′ flanking region of 30–40% of *CYP2A6*1* alleles results in genotype misclassification. In fact, no true *CYP2A6*3* alleles were found among 100 Spaniards and 96 Chinese subjects. In one Spanish poor metaboliser of the CYP2A6 probe drug coumarin, we found two novel defective alleles. One, *CYP2A6*5*, encoded an unstable enzyme having a G479L substitution and the other was found to carry a novel type of *CYP2A6* gene deletion (*CYP2A6*4D*). The results imply the presence of numerous defective as well as active *CYP2A6* alleles as a consequence of *CYP2A6*/*CYP2A7* gene conversion events. We conclude that molecular epidemiological studies concerning CYP2A6 require validated genotyping methods for accurate detection of all known defective *CYP2A6* alleles.

Key words: Cytochrome P450; Cytochrome P450 2A6; Coumarin; Nicotine; Cotinine; Lung cancer

1. Introduction

The cytochromes P450 are a superfamily of enzymes involved in the metabolism of numerous exogenous and endogenous substances including drugs, environmental chemicals, steroid hormones and bile acids. Many of the genes that encode the drug metabolising P450s are to a high extent polymorphic, thereby causing pronounced interindividual variability in the metabolism of many clinically used drugs [1].

Cytochrome P450 2A6 (CYP2A6) is a hepatic P450 which metabolises certain pharmaceutical agents, e.g. coumarin, (+)-*cis*-3,5-dimethyl-2-(3-pyridyl) thiozolidin-4-one hydrochloride (SM-12502), methoxyflurane, halothane, losigamone, letrozole, valproic acid and disulfiram. The enzyme can also activate a number of precarcinogens, including 4-methylnitrosamino-1-(3-pyridyl)-1-butanone (NNK), *N*-nitrosodiethylamine, 1,3-butadiene and 2,6-dichlorobenzonitrile (see [2] for references). In humans, the major pathway for nicotine metabolism consists of its *C*-oxidation to cotinine followed by cotinine 3′-hydroxylation [3], with CYP2A6 being the major enzyme that catalyses both of these reactions [4–8].

Phenotyping studies done with the probe drug coumarin have revealed pronounced interindividual variability in CYP2A6 activity in vivo [9–12]. Similar results have been obtained when CYP2A6 levels and activity were determined in human liver microsomes [13–15]. Despite the great interindividual variation observed, very few individuals lacking CYP2A6 activity have been found in Caucasians [2,16]. The frequency of poor metabolisers (PMs) in European and Middle East populations is $\leq 1\%$, whereas it is much higher in Asian populations.

Genetic polymorphisms of the *CYP2A6* gene can probably explain this interindividual variability to a great extent. Two types of defective alleles have been described thus far. The *CYP2A6*2*[1] allele encodes an enzyme with a L160H substitution that does not incorporate haem and is therefore inactive [17]. We and others [18,19] have also described some inactive alleles common in Asian populations where parts of or the whole *CYP2A6* gene has been deleted (*CYP2A6*4A*, **4B* and **4C*). In addition, *CYP2A6*3*, described as a hybrid allele generated by multiple gene conversions with the inactive *CYP2A7* gene, has been proposed to be inactive [20] but this has not yet been demonstrated.

Pianezza and co-workers [21] hypothesised that the *CYP2A6* genotype would be of importance for an individua's smoking behaviour. They showed that a lower number of individuals carrying *CYP2A6*2* and *CYP2A6*3* alleles, detected by the method described by Fernandez-Salguero et al, [20], were found in a tobacco-dependent group as compared to a never tobacco-dependent group and that subjects carrying these alleles smoked fewer cigarettes. A recent study supports an important role for the *CYP2A6* genotype in nicotine metabolism in vivo, whereby individuals homozygous for a *CYP2A6* gene deletion displayed only 15% of the levels of cotinine in urine as compared to subjects carrying at least one active gene when they smoked the same number of cigarettes [8]. Furthermore, Japanese subjects with a *CYP2A6* gene deletion were shown to be at reduced risk for lung cancer in a preliminary study [22].

*Corresponding author. Fax: (46) (8) 33 73 27.
E-mail: mikael.oscarson@imm.ki.se

Abbreviations: CYP or P450, cytochrome P450; wt, wild type; PM, poor metaboliser; PCR, polymerase chain reaction

[1] We have used the nomenclature system for *CYP2A6* alleles recommended by the Human Cytochrome P450 Allele Nomenclature Committee, see http://www.imm.ki.se/CYPalleles/

PII: S0014-5793(99)01364-2

We have previously reported that the genotyping method used for *CYP2A6*2* in the studies by Pianezza and co-workers [21] gave erroneous CYP2A6 phenotype predictions and developed an improved method for the detection of the *CYP2A6*2* allele [2]. In this report we have extended this method for the accurate detection of the *CYP2A6*3* allele and show that it is not present among the Caucasian and Chinese populations investigated. In addition, we have characterised new *CYP2A6* alleles whose presence might have important implications for the analysis of e.g. smoking behaviour and risk of lung cancer in relation to the *CYP2A6* genotype.

2. Materials and methods

2.1. Subjects

Genomic DNA was obtained from Spanish subject H83 who was previously scored as homozygous for the *CYP2A6*3* allele [23]. Genomic DNA was also obtained from 100 Spaniards and 96 Chinese individuals involved in earlier CYP2D6 genotype-phenotype correlation studies [24,25]. The study was approved by the ethical committees at Karolinska Institutet, the University of Oulu and the University Hospital Infanta Cristina in Badajoz.

2.2. CYP2A6 genotyping

Genotyping for the *CYP2A6*2* allele was performed as previously described [2] with the addition of 0.25 μM of primers 2A6*3wt and 2A6*3mut, respectively (see Table 1) in PCR II for simultaneous detection of the *CYP2A6*3* allele. The corresponding region from exon 1 to exon 4 of the *CYP2A7* gene was used as a positive control for the *CYP2A6*3* allele.

A two-step PCR method was used to detect the *CYP2A6*5* and *CYP2A6*1B* alleles. In the first PCR reaction (PCR I), a region from exon 8 to the 3′ flanking region of the *CYP2A6* gene was specifically amplified. This PCR product was subsequently used as a template in the second allele-specific PCR reaction (PCR II). PCR I was done using primers 2A6ex8F and 2A6R1 (see Table 1). The reaction mixture contained approximately 50 ng genomic DNA, 0.25 μM of each primer, 0.2 mM of each dNTP, 1.0 mM $MgCl_2$, 0.6 U of Taq polymerase (Life Technologies, Rockville, MD, USA) and was carried out in 1×PCR Buffer (Life Technologies) in a total volume of 25 μl using a Perkin Elmer GeneAmp PCR system 2400/9700. Initial denaturation was performed at 95°C for 1 min, followed by 35 cycles each consisting of denaturation at 95°C for 15 s, annealing at 56°C for 20 s, and extension at 72°C for 1.5 min, followed by a final extension at 72°C for 7 min.

The mixture for the *CYP2A6*5* allele-specific reaction (PCR II) contained 0.8 μl PCR I product, 0.25 μM of primer 2A6*5wt or primer 2A6*5mut, 0.25 μM of primer 2A6R2 (see Table 1), 0.1 mM of each dNTP, 1.0 mM $MgCl_2$, 0.3 U of Taq polymerase, and was done in 1× PCR buffer (Life Technologies) in a total volume of 25 μl. The amplification was performed by initial denaturation at 95°C for 1 min, followed by 15 cycles each involving denaturation at 95°C for 15 s, annealing at 54°C for 20 s, and extension at 72°C for 45 s. Ten μl of the PCR products were analysed on a 2% agarose gel (Life Technologies) stained with ethidium bromide. A similar method was used for detection of the *CYP2A6*1B* allele, except that the primers 2A6*1Bwt and 2A6*1Bmut were used in PCR II.

2.3. Southern blot analysis

Southern Blot analysis with the restriction enzyme *Eco*RI was carried out as previously reported [18].

2.4. Long PCR and DNA sequencing

Segments of the 5′ flanking regions of the *CYP2A6* and *CYP2A7* genes were sequenced using cosmids 19296 and 19019 as templates [26] (kindly supplied by Dr Linda K. Ashworth, LLNL Human Genome Center, Livermore, CA, USA). The *CYP2A6* and *CYP2A7* genes, as well as the *CYP2A7/CYP2A6* hybrid, were amplified from individual H83 using a GeneAmp XL PCR Kit (Perkin-Elmer, Norwalk, CT, USA) and different combinations of primers 2A6F1, 2A7F1, 2A6R1 and 2A7R1 (see Table 1). All exons, exon-intron junctions and 5′ and 3′ flanking regions were sequenced with the ABI Prism BigDye terminator cycle sequencing kit and analysed on an ABI Prism 377 DNA sequencer.

2.5. Site-directed mutagenesis and subcloning of CYP2A6 cDNA into the pYe60 yeast expression vector

A wt CYP2A6 cDNA [27] was used as a template in a PCR amplification with the eLONGase system (Life Technologies) using different combinations of the primers V60-2A6F, V60-2A6F-Δ5, V60-2A6Rwt and V60-2A6Rmut (see Table 1) to produce constructs with the two potential translation initiation sites with or without the 1436G→T mutation in exon 9. The forward primers also introduced an AAA in front of the translation initiation site which has previously been shown to increase expression of CYP2D6 [28]. The PCR fragments were subsequently cloned into the *Bam*HI and *Kpn*I sites of the pYe60 (V60) yeast expression vector [29] and sequenced to exclude any potential PCR artefacts.

2.6. Expression in yeast cells

Saccharomyces cerevisiae strain *W(R)* [30], which has been genetically modified to overexpress the yeast cytochrome P450 reductase gene (Yred), was transfected with the V60-2A6 plasmids described above. Expression of CYP2A6 was carried out as previously described [31,32] except that yeast cell microsomes were prepared by ultracentrifugation (100 000×*g* for 60 min). Microsomal protein concentrations and total cytochrome P450 content were determined with previously described methods [33,34].

2.7. Assay of CYP2A6 apoprotein levels and coumarin 7-hydroxylase activity

To quantify apoprotein levels, 5 μg of microsomal protein was subjected to sodium dodecyl sulphate gel electrophoresis using a 10% polyacrylamide gel. The proteins were transferred to a nitrocellulose filter (Bio-Rad, Hemel Hempstead, UK), incubated with the monoclonal CYP2A6 antibody MAB-2A6 (Gentest, Woburn, MA, USA) and then with a secondary horseradish peroxidase-linked goat anti-mouse antibody (Dako AS, Glostrup, Denmark). The enhanced chemiluminescence method (Pierce, Rockford, IL, USA) was used to visualise the CYP2A6 proteins.

Coumarin 7-hydroxylase activity was determined in reaction mix-

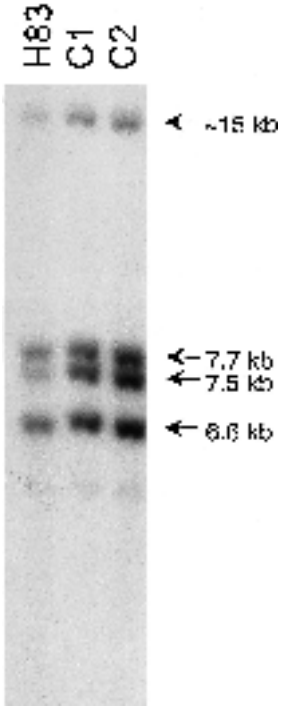

Fig. 1. Southern blot analysis of the *CYP2A* locus in genomic DNA from individual H83 using the restriction endonuclease *Eco*RI. C1 and C2 represent genomic DNA from control individuals carrying two copies of the *CYP2A6* gene.

Table 1
Primers used in the present study

Primer	Sequence
V60-2A6F	5′-gatcagatctaaaatgctggcctcagggatgcttctggtggcc-3′
V60-2A6F-Δ5	5′-gatcagatctaaaatgcttctggtggccttgctg-3′
V60-2A6Rwt	5′-gatcggtacctcagcggggcaggaagctcatg-3′
V60-2A6Rmut	5′-gatcggtacctcagcggggcaggaagctcatggtgtagtttcgtgggatcgtggcaaagaccacgtg-3′
2A6ex1[a]	5′-gctgaacacagagcagatgtaca-3′
2A6ex4R[a]	5′-ggaggttgacgtgaactggaaga-3′
2A6*2wt[a]	5′-ctcatcgacgccct-3′
2A6*2mut[a]	5′-ctcatcgacgccca-3′
2A6*3wt	5′-gctccggcgcttct-3′
2A6*3mut	5′-gctcctgcgctttg-3′
E3R[b]	5′-tcgtcctgggtgttttccttc-3′
2A6F1	5′-tggctgtgtcccaagctaggca-3′
2A7F1	5′-tggctgtgtcccaagctaggtg-3′
2A6R1[c]	5′-gcacttatgttttgtgagacatcagagacaa-3′
2A7R1	5′-gcacttatgttttgtgagacatcagatagag-3′
2A6R2[c]	5′-aaaatgggcatgaacgccc-3′
2A6ex8F	5′-ccagcacttcctgaatgag-3′
2A6*5wt	5′-ccccaaacacgtggg-3′
2A6*5mut	5′-ccccaaacacgtggt-3′
2A6*1Bwt	5′-actgggggcaggatggc-3′
2A6*1Bmut	5′-aatgggggggaagatgcg-3′

Restriction enzyme sites used for cloning are underlined.
[a]From Oscarson et al. [2].
[b]From Fernandez-Salguero et al. [20].
[c]From Oscarson et al. [18].

tures consisting of 100 mM Tris-HCl pH 7.4, 250 μg of yeast microsomes and 100 μM of coumarin in a total volume of 500 μl. The reactions were initiated by addition of 0.5 μmol of NADPH and terminated with 500 μl 6% trichloroacetic acid. The samples were subsequently centrifuged at 15 000×*g* for 10 min and the amount of 7-OH-coumarin formed was measured using a fluorometric method [35].

3. Results

3.1. *CYP2A6*3* genotyping

In order to accurately evaluate the presence of the *CYP2A6*3* allele and to avoid possible erroneous PCR amplification, we developed a modified method for its detection using the primers in exon 1 and exon 4 that were previously designed for *CYP2A6*-specific amplification [2] with allele-specific PCR primers that simultaneously detect the *CYP2A6*2* and *CYP2A6*3* alleles. Using this method, we could not find any *CYP2A6*3* allele in the two populations studied, namely 100 Spaniards and 96 Chinese individuals (Table 2). This is in contrast to the 2–28% previously reported, indicating that the great majority of the individuals who were found to be positive for the *CYP2A6*3* allele with the method described by Fernandez-Salguero and co-workers [20] are in fact false positives (see below).

3.2. Southern blot and sequencing of genomic DNA from individual H83

To clarify the true origin of the alleles which were classified

Table 2
CYP2A6 genotypes and allele frequencies in a Spanish and a Chinese population determined by allele-specific PCR methods

		Population	
		Spaniards ($n=100$)	Chinese ($n=96$)
Genotypes (%)			
*CYP2A6*1A/CYP2A6*1A*		48.0	19.8
*CYP2A6*1A/CYP2A6*1B*		33.0	36.5
*CYP2A6*1B/CYP2A6*1B*		12.0	14.6
*CYP2A6*1A/CYP2A6*2*		4.0	0
*CYP2A6*1B/CYP2A6*2*		2.0	0
*CYP2A6*1A/CYP2A6*4*		0	10.4
*CYP2A6*1B/CYP2A6*4*		1.0	13.5
*CYP2A6*4/CYP2A6*4*		0	3.1
*CYP2A6*1B/CYP2A6*5*		0	2.1
Alleles (%)	Functional consequence		
*CYP2A6*1A*	wt	66.5	43.2
*CYP2A6*1B*	gene conversion in the 3′ flanking region	30.0	40.6
*CYP2A6*2*	L160H	3.0	0
*CYP2A6*3*	multiple *CYP2A6*/*CYP2A7* gene conversions	0	0
*CYP2A6*4*	CYP2A6 gene deletion	0.5	15.1
*CYP2A6*5*	G479L	0	1.0

Data compiled from this and previous studies [2,18].

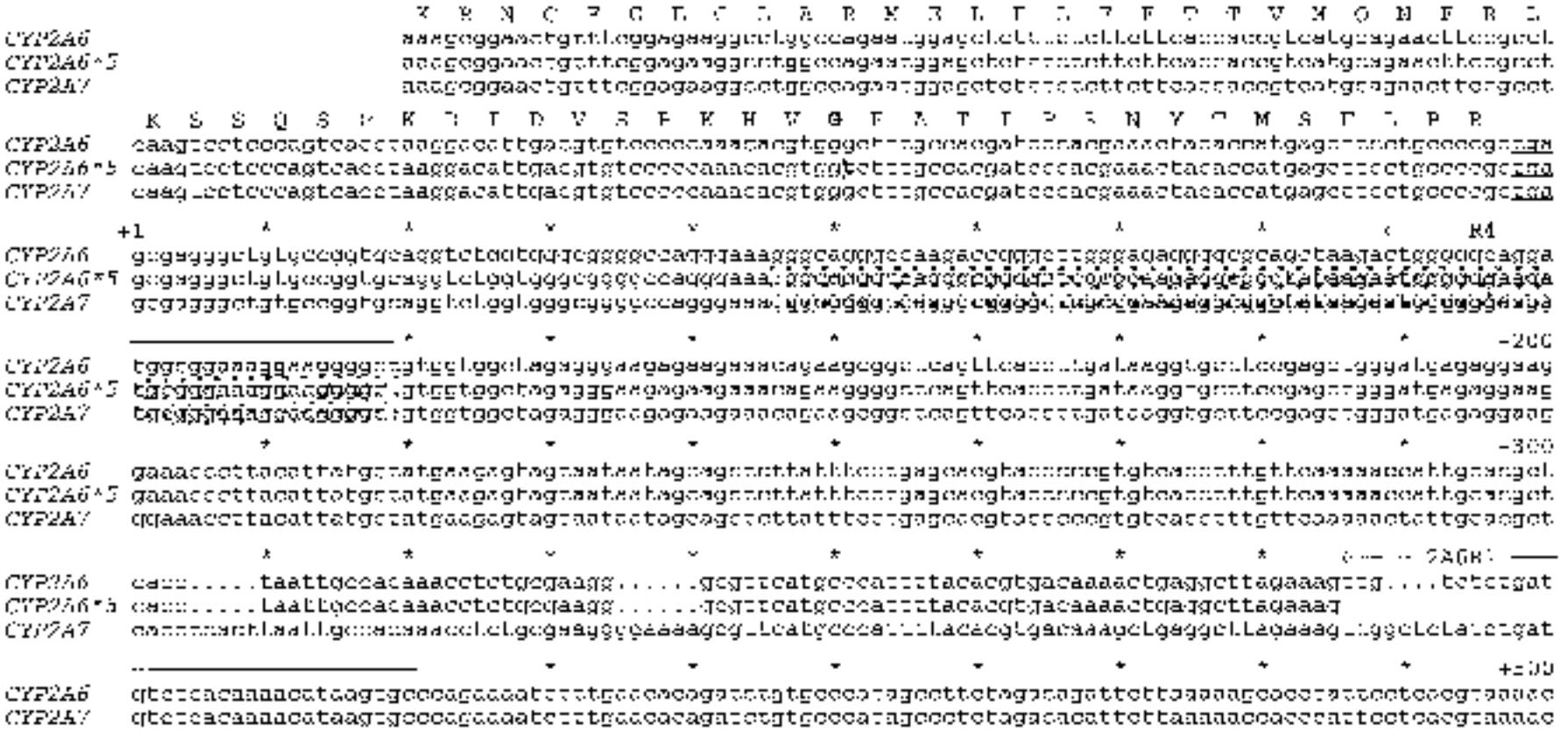

Fig. 2. DNA sequence of exon 9 and the 3′ flanking regions of the *CYP2A6* and *CYP2A7* genes and the *CYP2A6*5* allele. The light grey shaded box represents the gene conversion with the *CYP2A7* gene. The 1436G→T mutation in exon 9 is also indicated with a shaded box. Arrows above the sequences indicate the primers used for amplification. Numbering of the sequence begins with the first base pair following the stop codon (underlined).

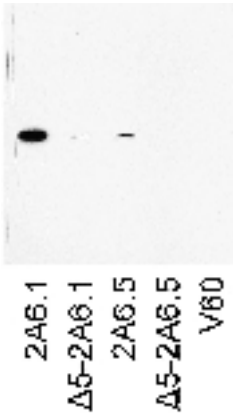

Fig. 3. Western blot analysis of the CYP2A6 enzyme variants expressed in *S. cerevisiae*. Experimental conditions are described in Section 2 and the samples are labelled as described in Table 3, with V60 representing yeast transfected with an empty vector.

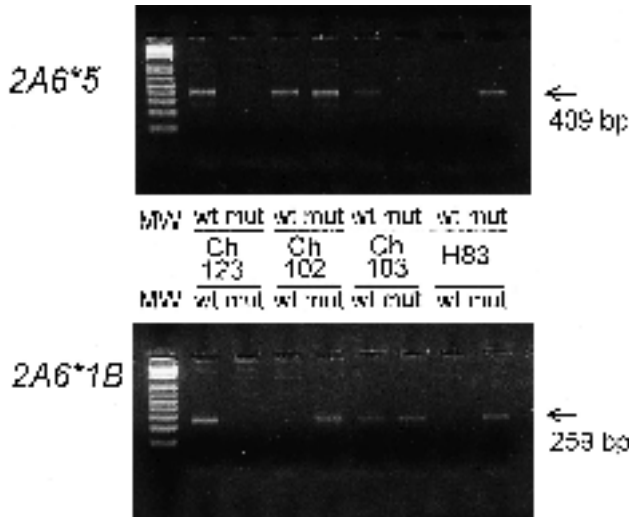

Fig. 4. Allele-specific PCR to detect the *CYP2A6*5* and *CYP2A6*1B* alleles ('PCR II'). MW, 1 kb plus DNA ladder (Life Technologies, Rockville, MD, USA). Genotypes of the subjects are: Ch123 (*CYP2A6*1A/*1A*), Ch102 (*CYP2A6*1B/*5*), Ch103 (*CYP2A6*1A/*1B*), H83 (*CYP2A6*4D/*5*).

as *CYP2A6*3*, we took advantage of an individual (H83) who was consistently scored as homozygous for the *CYP2A6*3* allele when assayed with the original genotyping method [20] but was homozygous for the *CYP2A6*1* allele with our genotyping method. Interestingly, phenotyping with the CYP2A6 probe drug coumarin demonstrated that this individual completely lacked CYP2A6 activity in vivo (data not shown).

Restriction fragment length polymorphism analysis of genomic DNA from individual H83 showed that one of the *CYP2A* genes was deleted (Fig. 1). Using different combinations of *CYP2A6*- and *CYP2A7*-specific primers in the 5′ and 3′ flanking regions, it was possible to amplify a *CYP2A6* gene, a *CYP2A7* gene as well as a *CYP2A7/CYP2A6* hybrid from this individual. The PCR products were subjected to DNA sequencing of all the exons and exon-intron junctions. The coding region of the *CYP2A6* allele was identical to the CYP2A6 cDNA sequence previously reported [17] except for a 1436G→T mutation in exon 9, yielding a G479L substitution. In addition, a gene conversion event with the *CYP2A7* gene was found in the 3′ flanking region (Fig. 2). This allele was designated *CYP2A6*5*.

Sequencing of the *CYP2A7/CYP2A6* hybrid revealed that it was consistent with a deletion of the *CYP2A6* gene, which most likely occurred through an unequal crossover event in a similar way as the *CYP2A6* gene deletion previously reported in Chinese individuals (*CYP2A6*4A*) [18]. Because exon 1 to exon 8 consists of the *CYP2A7* sequence and the 3′ flanking region has the *CYP2A6* sequence, we concluded that the crossover junction region is located in either intron 8 or exon 9, in contrast to the *CYP2A6*4A* allele where it is located in the 3′ flanking region. This new type of deletion was designated *CYP2A6*4D* and can also be detected using our PCR-based genotyping method for *CYP2A6* deletions [18] (data not shown).

3.3. Expression of CYP2A6*5 in S. cerevisiae

In order to evaluate the effect of the G479L substitution on the CYP2A6 enzyme, the 1436G→T mutation was introduced into a wt CYP2A6 cDNA. Furthermore, two different N-terminal variants were generated to correspond to the two potential translation initiation sites reported [17,27]. Heterologous expression of the different constructs in *S. cerevisiae* revealed that the first translation initiation site resulted in substantially higher expression levels in our system (Table 3 and Fig. 3). Introduction of the G479L substitution gave an enzyme variant (CYP2A6.5) which was only expressed at very low levels with a concomitant lack of coumarin 7-hydroxylation activity (Table 3 and Fig. 3), indicating that this substitution yields a much less stable enzyme that is rapidly degraded.

3.4. CYP2A6*1B and CYP2A6*5 genotyping

A PCR-based genotyping method was developed to detect the *CYP2A6*5* allele and was used to genotype two different populations (Fig. 4). It was determined that this allele was very rare in the Chinese and Spanish populations examined, with an allele frequency of 0–1% (Table 2).

Interestingly, the *CYP2A6*5* allele also contained a gene

Table 3
Characterisation of yeast microsomes from the *S. cerevisiae* P450 reductase overexpressing strain *W(R)* transfected with *CYP2A6*1* and *CYP2A6*5* alleles

Enzyme variant	P450 holoprotein level (pmol/mg microsomal protein)	Coumarin 7-OH activity (pmol/min/mg microsomal protein)
CYP2A6.1	44 ± 21	310 ± 140
Δ5-CYP2A6.1	ND	ND
CYP2A6.5	ND	ND
Δ5-CYP2A6.5	ND	ND

In enzyme variants labelled with Δ5, the second potential translation initiation site was used. The data represent mean ± S.D. of experiments from 2–6 batches of transfected cells. ND, not detectable.

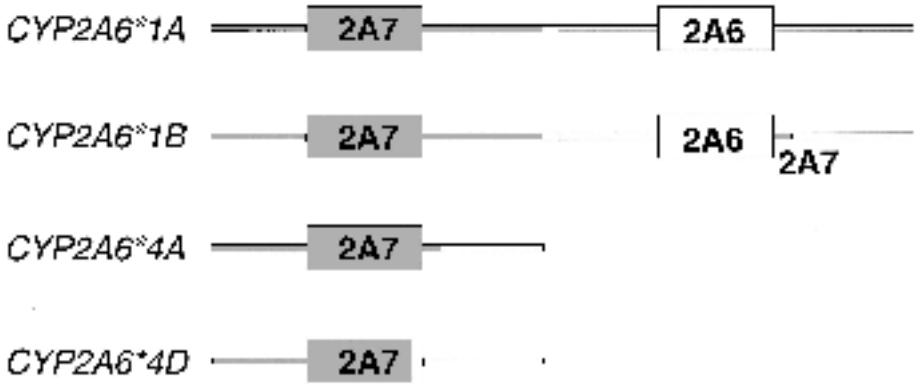

Fig. 5. Schematic diagram of the known *CYP2A6* alleles that have been created through various cross-over events. Shaded and unshaded regions represent *CYP2A7* and *CYP2A6* sequences, respectively.

conversion with the *CYP2A7* gene in the 3′ flanking region of the *CYP2A6* gene. This is the region where the reverse primer R4 is located in the original genotyping method [20] and results in a failure to amplify the *CYP2A6* gene. Genotyping for this gene conversion revealed an allele frequency of 30.0 and 40.6% in the Spanish and Chinese populations, respectively (Fig. 4 and Table 2). Its presence most likely explains the high frequency of the *CYP2A6*3* allele previously reported (see below). This allele, without the 1436G→T mutation, was denoted *CYP2A6*1B*.

4. Discussion

In this report we have identified and characterised two novel inactive *CYP2A6* alleles, the *CYP2A6*5* allele which has a 1436G→T mutation in exon 9 yielding a G479L substitution and *CYP2A6*4D*, a novel *CYP2A6* gene deletion variant. In addition, we found a common *CYP2A7* gene conversion in the 3′ flanking region of the *CYP2A6* gene. An overview of the known *CYP2A6* alleles that have been generated through different unequal crossover events is shown in Fig. 5.

The 1436G→T mutation is also found in one of the *CYP2A7* sequences reported (accession number U22029 [20]), therefore it is likely that this mutation has been transferred from the related *CYP2A7* gene through various cross-over events. This is in analogy with e.g. the *CYP21* and *CYP2D6* genes, where it appears that many of the inactivating mutations are continuously transferred from the corresponding pseudogenes *CYP21P* and *CYP2D7P* [36,37]. The G479 amino acid residue is located in substrate recognition site 6 [38] and is highly conserved in the CYP2A, CYP2B and CYP2F subfamilies. All of the 28 P450s in these subfamilies that have been reported to the SwissProt database, except CYP2A8 from golden hamster, have a glycine residue at the corresponding position. This indicates an important structural function for this residue, which is often the case for glycine residues, and consistent with our findings that showed the G479L substitution to yield a very unstable enzyme. We cannot, however, exclude that this mutation caused a decrease in mRNA stability or translation efficiency.

We have demonstrated that the commonly used *CYP2A6* genotyping method [20] gives erroneous results with respect to the *CYP2A6*3* allele, the predominant allele in many studies which is considered to be inactive. This method is based on initial amplification of the whole *CYP2A6* gene with the primers F4 and R4, followed by amplification of the diagnostic exon 3 and subsequent digestion with the restriction enzymes *Xcm*I and *Dde*I. In the second PCR reaction, 31 cycles of amplification are used, and this high number of cycles is enough for the amplification of carry-over genomic DNA in cases where the first PCR reaction failed for some reason. In individual H83, this method does not amplify the *CYP2A6*5* allele because of the gene conversion in the 3′ flanking region. Instead, the *CYP2A7*/*CYP2A6* hybrid is amplified from the *CYP2A6*4D* allele which has the *CYP2A6* sequence in the 3′ flanking region. Further analysis of exon 3 reveals only the *CYP2A7* sequence and the individual is therefore erroneously scored as homozygous for the *CYP2A6*3* allele.

Interestingly, genomic DNA from the 12–15% of the population that are homozygous for the *CYP2A6*1B* allele would not be amplified in the first PCR reaction using the original method. However, both the *CYP2A6* and *CYP2A7* genes would be amplified from the carry-over genomic DNA and these individuals would be erroneously scored as heterozygous for the *CYP2A6*3* allele (data not shown). This most likely explains the high *CYP2A6*3* allele frequency reported in many studies [20,23,39].

We conclude that it is very important that all genotyping methods are validated in large populations of phenotyped individuals before they are used in molecular epidemiological studies. This is especially important when complex genetic loci are analysed and the risk of coamplification of related genes exist. The method used by Pianezza and co-workers [21] detects a high frequency of the *CYP2A6*3* allele, whereas our more specific genotyping method revealed the complete absence of this allele. The negligible frequency of the *CYP2A6*3* allele has also been confirmed in two recent studies [8,40]. There is a growing need for accurate genotype-phenotype studies with respect to a true evaluation of the hypothesis raised by Pianezza and co-workers, and the genotyping methods and new alleles presented here will aid in further investigations in this field.

Acknowledgements: We are indebted to Drs Maria Luisa Bernal Ruiz, Blanca Sinues and Qun-Ying Yue for access to the genomic DNA samples, to Dr Linda K. Ashworth (LLNL Human Genome Center, Livermore, CA, USA) for providing the cosmids and to Dr Denis Pompon (Centre de Génétique Moléculaire du CNRS, Gif-sur-Yvette, France) for providing the *Saccharomyces cerevisiae* strain *W(R)* and

the pYeD60 vector. The study was supported by grants from the Swedish Society for Medical Research, Swedish Match and the Swedish Medical Research Council.

References

[1] Ingelman-Sundberg, M., Oscarson, M. and McLellan, R.A. (1999) Trends Pharmacol. Sci. 20, 342–349.
[2] Oscarson, M., Gullstén, H., Rautio, A., Bernal, M.L., Sinues, B., Dahl, M.-L., Stengård, J.H., Pelkonen, O., Raunio, H. and Ingelman-Sundberg, M. (1998) FEBS Lett. 438, 201–205.
[3] Benowitz, N.L. and Jacob, P. (1997) NIDA Res. Monogr. 173, 48–64.
[4] Nakajima, M., Yamamoto, T., Nunoya, K., Yokoi, T., Nagashima, K., Inoue, K., Funae, Y., Shimada, N., Kamataki, T. and Kuroiwa, Y. (1996) Drug Metab. Dispos. 24, 1212–1217.
[5] Messina, E.S., Tyndale, R.F. and Sellers, E.M. (1997) J. Pharmacol. Exp. Ther. 282, 1608–1614.
[6] Yamazaki, H., Inoue, K., Hashimoto, M. and Shimada, T. (1999) Arch. Toxicol. 73, 65–70.
[7] Nakajima, M., Yamamoto, T., Nunoya, K., Yokoi, T., Nagashima, K., Inoue, K., Funae, Y., Shimada, N., Kamataki, T. and Kuroiwa, Y. (1996) J. Pharmacol. Exp. Ther. 277, 1010–1015.
[8] Kitagawa, K., Kunugita, N., Katoh, T., Yang, M. and Kawamoto, T. (1999) Biochem. Biophys. Res. Commun. 262, 146–151.
[9] Iscan, M., Rostami, H., Iscan, M., Guray, T., Pelkonen, O. and Rautio, A. (1994) Eur. J. Clin. Pharmacol. 47, 315–318.
[10] Rautio, A., Kraul, H., Kojo, A., Salmela, E. and Pelkonen, O. (1992) Pharmacogenetics 2, 227–233.
[11] Cholerton, S., Idle, M.E., Vas, A., Gonzalez, F.J. and Idle, J.R. (1992) J. Chromatogr. 575, 325–330.
[12] Hadidi, H., Irshaid, Y., Vagbo, C.B., Brunsvik, A., Cholerton, S., Zahlsen, K. and Idle, J.R. (1998) Eur. J. Clin. Pharmacol. 54, 437–441.
[13] Shimada, T., Yamazaki, H. and Guengerich, F.P. (1996) Xenobiotica 26, 395–403.
[14] Yun, C.H., Shimada, T. and Guengerich, F.P. (1991) Mol. Pharmacol. 40, 679–685.
[15] Pearce, R., Greenway, D. and Parkinson, A. (1992) Arch. Biochem. Biophys. 298, 211–225.
[16] Hadidi, H., Zahlsen, K., Idle, J.R. and Cholerton, S. (1997) Food Chem. Toxicol. 35, 903–907.
[17] Yamano, S., Tatsuno, J. and Gonzalez, F.J. (1990) Biochemistry 29, 1322–1329.
[18] Oscarson, M., McLellan, R.A., Gullstén, H., Yue, Q.-Y., Lang, M.A., Bernal, M.L., Sinues, B., Hirvonen, A., Raunio, H., Pelkonen, O. and Ingelman-Sundberg, M. (1999) FEBS Lett. 448, 105–110.
[19] Nunoya, K.I., Yokoi, T., Kimura, K., Kainuma, T., Satoh, K., Kinoshita, M. and Kamataki, T. (1999) J. Pharmacol. Exp. Ther. 289, 437–442.
[20] Fernandez-Salguero, P., Hoffman, S.M., Cholerton, S., Mohrenweiser, H., Raunio, H., Rautio, A., Pelkonen, O., Huang, J.D., Evans, W.E., Idle, J.R. and Gonzalez, F.J. (1995) Am. J. Hum. Genet. 57, 651–660.
[21] Pianezza, M.L., Sellers, E.M. and Tyndale, R.F. (1998) Nature 393, 750.
[22] Miyamoto, M., Umetsu, Y., Dosaka-Akita, H., Sawamura, Y., Yokota, J., Kunitoh, H., Nemoto, N., Sato, K., Ariyoshi, N. and Kamataki, T. (1999) Biochem. Biophys. Res. Commun. 261, 658–660.
[23] Gullstén, H., Agúndez, J.A., Benítez, J., Laara, E., Ladero, J.M., Diaz-Rubio, M., Fernandez-Salguero, P., Gonzalez, F., Rautio, A., Pelkonen, O. and Raunio, H. (1997) Pharmacogenetics 7, 247–250.
[24] Bernal Ruiz, M.L., Sinues, B., Johansson, I., McLellan, R.A., Dahl, M.-L., Ingelman-Sundberg, M. and Bertilsson, L. (1999) Pharmacogenetics (in press).
[25] Johansson, I., Yue, Q.Y., Dahl, M.L., Heim, M., Säwe, J., Bertilsson, L., Meyer, U.A., Sjöqvist, F. and Ingelman-Sundberg, M. (1991) Eur. J. Clin. Pharmacol. 40, 553–556.
[26] Hoffman, S.M., Fernandez-Salguero, P., Gonzalez, F.J. and Mohrenweiser, H.W. (1995) J. Mol. Evol. 41, 894–900.
[27] Salonpää, P., Hakkola, J., Pasanen, M., Pelkonen, O., Vähäkangas, K., Battula, N., Nouso, K. and Raunio, H. (1993) Eur. J. Pharmacol. 248, 95–102.
[28] Krynetski, E.Y., Drutsa, V.L., Kovaleva, I.E. and Luzikov, V.N. (1995) Pharmacogenetics 5, 103–109.
[29] Urban, P., Cullin, C. and Pompon, D. (1990) Biochimie 72, 463–472.
[30] Truan, G., Cullin, C., Reisdorf, P., Urban, P. and Pompon, D. (1993) Gene 125, 49–55.
[31] Bellamine, A., Gautier, J.C., Urban, P. and Pompon, D. (1994) Eur. J. Biochem. 225, 1005–1013.
[32] Oscarson, M., Hidestrand, M., Johansson, I. and Ingelman-Sundberg, M. (1997) Mol. Pharmacol. 52, 1034–1040.
[33] Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) J. Biol. Chem. 193, 265–275.
[34] Omura, T. and Sato, R. (1964) J. Biol. Chem. 239, 2370–2378.
[35] Aitio, A. (1978) Anal. Biochem. 85, 488–491.
[36] Wedell, A. (1998) Acta Paediatr. 87, 159–164.
[37] Daly, A.K., Brockmöller, J., Broly, F., Eichelbaum, M., Evans, W.E., Gonzalez, F.J., Huang, J.D., Idle, J.R., Ingelman-Sundberg, M., Ishizaki, T., Jacqz-Aigrain, E., Meyer, U.A., Nebert, D.W., Steen, V.M., Wolf, C.R. and Zanger, U.M. (1996) Pharmacogenetics 6, 193–201.
[38] Gotoh, O. (1992) J. Biol. Chem. 267, 83–90.
[39] London, S.J., Idle, J.R., Daly, A.K. and Coetzee, G.A. (1999) Lancet 353, 898–899.
[40] Chen, G.F., Tang, Y.M., Green, B., Lin, D.X., Guengerich, F.P., Daly, A.K., Caporaso, N.E. and Kadlubar, F.F. (1999) Pharmacogenetics 9, 327–332.