Novel human *CYP2A6* alleles confound gene deletion analysis

Miki Nakajima[a], Ryoko Yoshida[a], Tatsuki Fukami[a], Howard L. McLeod[b], Tsuyoshi Yokoi[a,*]

[a]*Drug Metabolism and Toxicology, Division of Pharmaceutical Sciences, Graduate School of Medical Science, Kanazawa University, Kakuma-machi, Kanazawa 920-1192, Japan*
[b]*Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110-1093, USA*

Received 20 February 2004; revised 17 May 2004; accepted 18 May 2004

Available online 7 June 2004

Edited by Takashi Gojobori

Abstract **Cytochrome P450 (CYP) 2A6 metabolizes a number of drugs and a variety of procarcinogens. CYP2A6 also catalyzes nicotine *C*-oxidation leading to cotinine formation, a major metabolic pathway of nicotine in humans. There are genetic polymorphisms in the human *CYP2A6* gene and a relationship between the *CYP2A6* genotype and smoking habits as well as the incidence of lung cancer has been indicated. *CYP2A6*4* alleles are the whole deleted type and are completely deficient in the enzymatic activity. An unequal crossover junction is located in the 3′-flanking region in the *CYP2A6*4A* allele, whereas the junction is located in either intron 8 or exon 9 in the *CYP2A6*4D* allele. In the present study, a novel genotyping method to distinguish between two different whole deleted alleles of *CYP2A6*4A* and *CYP2A6*4D* was established. In the process, two novel alleles, *CYP2A6*1F* and *CYP2A6*1G*, were found. The *CYP2A6*1F* has a single nucleotide polymorphism (SNP) of C5717T in exon 8, and the *CYP2A6*1G* has two SNPs, C5717T in exon 8 and A5825G in intron 8. The SNP of C5717T corresponds to C1224T on the cDNA sequence and is a synonymous mutation. Since the *CYP2A6*1F* produces a recognition site of the restriction enzymes that is the same as *CYP2A6*4D*, the presence of the *CYP2A6*1F* allele could cause a mistyping as the *CYP2A6*4D* allele. According to an improved genotyping method, the allele frequencies of *CYP2A6*4A*, *CYP2A6*4D*, *CYP2A6*1F*, and *CYP2A6*1G* in 165 Caucasians were 3.0%, 0%, 1.8%, and 1.2%, respectively. The allele frequencies of *CYP2A6*4A, CYP2A6*4D, CYP2A6*1F*, and *CYP2A6*1G* in 94 African-Americans were 0%, 0.5%, 0%, and 13.3%, respectively. This is the first report of a method that can distinguish between *CYP2A6*4A*, *CYP2A6*4D*, and *CYP2A6*1F* which could otherwise cause a mistyping as *CYP2A6*4D*.**

© 2004 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Cytochrome P450; Genetic polymorphism; Interindividual difference; Nicotine metabolism; Coumarin 7-hydroxylation

1. Introduction

Cytochrome P450 (CYP) is a superfamily of hemoproteins, many of which can metabolize xenobiotics such as procarcinogens, drugs, and environmental pollutants. CYP2A6 is a major hepatic member of the family in humans that metabolizes pharmaceutical agents such as coumarin, methoxyflurane, halothane, losigamone, letrozole, valproic acid, disulfiram, and fadrozole, and activates some procarcinogens such as 4-methylnitrosoamino-1-(3-pyridyl)-1-butanone and *N*-nitrosodiethylamine [1]. Especially, CYP2A6 is a major metabolic enzyme of nicotine. Nicotine is metabolized to cotinine by CYP2A6 [2], and cotinine is further metabolized to *trans*-3′-hydroxycotinine by CYP2A6 [3]. In the *CYP2A6* gene, several mutated alleles have been reported. The *CYP2A6*1A* is a wild type of the *CYP2A6* gene. The *CYP2A6*1B* allele has a gene conversion with *CYP2A7* in the 3′-untranslated region [4]. The *CYP2A6*3* allele has gene conversions with *CYP2A7* in exons 3, 6, and 8 [5]. CYP2A7 is catalytically inactive. The *CYP2A6*2*, *CYP2A6*5*, *CYP2A6*6*, *CYP2A6*7*, *CYP2A6*8*, and *CYP2A6*11* alleles have single amino acid substitutions of L160H, G479V, R128Q, I471T, R485L, and S224P, respectively [4,6–9]. The *CYP2A6*9* allele has a point mutation in the TATA box (T-48G) [10]. The *CYP2A6*10* allele has two simultaneous amino acid substitutions of *CYP2A6*7* and *CYP2A6*8* [11,12]. The *CYP2A6*4* allele deletes the whole *CYP2A6* gene [13–17]. The *CYP2A6*1X2* allele has a duplication of the *CYP2A6* gene [18]. This allele is considered to be the reciprocal product of the *CYP2A6*4* allele after an unequal crossover event between the 3′-flanking regions of the *CYP2A6* and *CYP2A7* genes. Recently, Oscarson et al. [19] reported a novel *CYP2A7*/*CYP2A6* hybrid allele (*CYP2A6*12*) that carries an unequal crossover in intron 2. Kiyotani et al. [20] reported additional alleles possessing a single amino acid substitution of G5R (*CYP2A6*13*), S29N (*CYP2A6*14*), K194E (*CYP2A6*15*), and R203S (*CYP2A6*16*).

In our previous studies [21–25], we investigated the relationship between the large interindividual differences in nicotine metabolism and the genetic polymorphisms of *CYP2A6*. We found that nicotine metabolism is impaired in the homozygotes of either *CYP2A6*4*, *CYP2A6*7* and *CYP2A6*10* in Japanese and Koreans [21–24]. Furthermore, we found that the *CYP2A6*9* allele causes a decrease in nicotine metabolism [25]. In Caucasians, it has been reported that homozygotes of the *CYP2A6*2* allele were deficient in nicotine metabolism [26]. Since the *CYP2A6*4* allele is a whole deletion type, the enzymatic activity is completely deficient. There are four types of *CYP2A6*4* alleles, i.e., *CYP2A6*4A*, *CYP2A6*4B*, *CYP2A6*4C*, and *CYP2A6*4D* (Fig. 1). The *CYP2A6*4A* allele lacks 3′-untranslated region of the *CYP2A7* gene and the entire *CYP2A6* gene and an unequal crossover junction is located in the 3′-untranslated region [4,13,14]. The *CYP2A6*4B* allele lacks the entire *CYP2A6* gene, whereas *CYP2A7* gene is normal [17]. The *CYP2A6*4C* allele is recognized to be the same as *CYP2A6*4A* [16]. For the *CYP2A6*4D* allele, an

* Corresponding author. Fax: +81-76-234-4407.
E-mail address: tyokoi@kenroku.kanazawa-u.ac.jp (T. Yokoi).

0014-5793/$22.00 © 2004 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2004.05.053

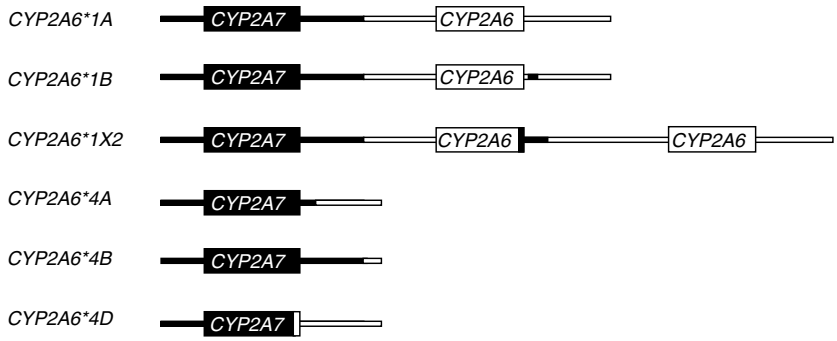

Fig. 1. Schematic diagram of the various *CYP2A6* alleles that have been created through crossover events. Closed and open regions represent *CYP2A7* and *CYP2A6* genes, respectively. *CYP2A7* is located 25 kb upstream of the *CYP2A6* gene.

unequal crossover junction is located in either intron 8 or exon 9 [15]. Oscarson et al. [15] found one Spanish subject possessing the *CYP2A6*4D* allele. However, the allele frequency is unknown since it was impossible to distinguish between the *CYP2A6*4A* or *CYP2A6*4D* alleles with the PCR method they used. In the present study, we developed a genotyping method to distinguish between the *CYP2A6*4A* and *CYP2A6*4D* alleles. In the process, two novel alleles possessing synonymous single nucleotide polymorphisms (SNPs) were found, and one of the alleles could be mistyped as the *CYP2A6*4D* allele. After the improvement of the genotyping method, the allele frequencies of the *CYP2A6*4A*, *CYP2A6*4D*, and two novel alleles of *CYP2A6*1F* and *CYP2A6*1G* in Caucasians and African-Americans were investigated.

2. Materials and methods

2.1. Chemicals and regents

Taq DNA polymerase was obtained from Greiner Japan (Tokyo, Japan). Restriction enzymes were purchased from Takara (Kyoto, Japan). Primers were commercially synthesized at Hokkaido System Sciences (Sapporo, Japan). All other chemicals and solvents were of the highest grade commercially available.

2.2. Genotyping of the CYP2A6*1A, CYP2A6*1B, CYP2A6*4A, and CYP2A6*4D alleles

This study was approved by the Human Studies Committee of Washington University School of Medicine (St. Louis, MO). One-hundred sixty five Caucasian (63 male and 102 female) and 94 African-American (26 male and 68 female) subjects were recruited. Written informed consent was obtained from all subjects. Blood samples were collected from a cubital vein. Genomic DNA was extracted from peripheral lymphocytes using a Puregene DNA isolation kit (Gentra Systems, Minneapolis, MN). The genotyping of *CYP2A6*1A*, *CYP2A6*1B*, *CYP2A6*4A*, and *CYP2A6*4D* was performed as follows: The 2Aex7F (5′-GRCCAAGATGCCCTACATG-3′) [15] and 2A6R2 (5′-AAAATGGGCATGAACGCCC-3′) [15] primers were used. The 2Aex7F primer could anneal the corresponding region of *CYP2A7* as well as that of *CYP2A6* The 2A6R2 primer could anneal only *CYP2A6*. Genomic DNA samples (0.5 μg) were added to the PCR mixtures (25 μl) consisting of 1× PCR buffer [67 mM Tris–HCl buffer (pH 8.8), 16.6 mM $(NH_4)_2SO_4$, 0.45% Triton X-100, 0.02% gelatin], 1.5 mM $MgCl_2$, 0.4 μM of each primer, 250 μM dNTPs, and 1 U of *Taq* DNA polymerase. After an initial denaturation at 95 °C for 1 minute, amplification was performed by denaturation at 95 °C for 15 s, annealing at 60 °C for 20 s, and extension at 72 °C for 3 min for 35 cycles, followed by a final extension at 72 °C for 7 min. The PCR product was double-digested with *Eco*81I and *Acc*II restriction enzymes. The digestion patterns were determined by electrophoresis in a 1.5% agarose gel. The schematic restriction fragment length polymorphism (RFLP) patterns are shown in Fig. 2B.

2.3. DNA sequencing

The PCR product with the primers of 2Aex7F and 2A6R2 was subcloned into pT7Blue T-vector (Novagen, Madison, WI). The plasmid DNA was purified by Qiagen Plasmid Midi kit (Qiagen, Valencia, CA) and submitted to DNA sequencing using a Thermo Sequenase Cy5.5 Dye Terminator Cycle Sequencing kit (Amersham Pharmacia Biotech, Buckinghamshire, UK) with T7F primer (Amersham Pharmacia Biotech). DNA sequences were analyzed on a Long-Read Tower DNA sequencer (Amersham Pharmacia Biotech). To determine the nucleotide sequences of the novel alleles of *CYP2A6*1F* and *CYP2A6*1G*, DNA sequences of all exons and exon–intron junctions were analyzed with direct sequence analyses using genomic DNA from a heterozygote of *CYP2A6*1A*/*CYP2A6*1F* and a homozygote of *CYP2A6*1G* allele.

2.4. Improved genotyping method for the CYP2A6*1A, CYP2A6*1B, CYP2A6*4A, CYP2A6*4D, and two novel alleles

To distinguish between the *CYP2A6*4D* and a novel allele, *CYP2A6*1F*, PCR-RFLP analysis was improved as follows: The 2Aint7F (5′-TTTGTGTCAGGAGAATCAAAC-3′) and 2A6R2 [15] primers were used. The 2Aint7F primer could anneal the corresponding region of *CYP2A7* as well as that of *CYP2A6*. Genomic DNA samples (0.5 μg) were added to the PCR mixtures (25 μl) consisting of 1× PCR buffer, 1.5 mM $MgCl_2$, 0.4 μM of each primer, 250 μM dNTPs, and 1 U of *Taq* DNA polymerase. After an initial denaturation at 94 °C for 3 min, amplification was performed by denaturation at 94 °C for 30 s, annealing at 53 °C for 30 s, and extension at 72 °C for 2 min for 30 cycles, followed by a final extension at 72 °C for 5 min. The PCR product was triple-digested with *Eco*81I, *Acc*II, and *Stu*I restriction enzymes. The digestion patterns were determined by electrophoresis in a 2% agarose gel. The schematic RFLP patterns are shown in Fig. 3B.

2.5. Distinction between CYP2A6*1A/CYP2A6*4A and CYP2A6*1B/CYP2A6*4D

In the genotyping method described above, the RFLP pattern of *CYP2A6*1A/CYP2A6*4A* and *CYP2A6*1B/CYP2A6*4D* are the same. To distinguish between these two genotypes, PCR was performed as follows: The 2Aint7F and 2A6UTR-RV (5′-AGTCTTAGCTGCGCCCCTC-3′) primers were used. The 2A6UTR-RV primer could specifically anneal the *CYP2A6*. Genomic DNA samples (0.5 μg) were added to the PCR mixtures (25 μl) consisting of 1× PCR buffer, 1.5 mM $MgCl_2$, 0.4 μM of each primer, 250 μM dNTPs, and 1 U of *Taq* DNA polymerase. After an initial denaturation at 94 °C for 3 min, amplification was performed by denaturation at 94 °C for 20 s, annealing at 54 °C for 20 s, and extension at 72 °C for 1.5 min for 35 cycles, followed by a final extension at 72 °C for 5 min. In the subjects with *CYP2A6*1A/CYP2A6*4A* and *CYP2A6*1B/CYP2A6*4D*, the *CYP2A6*1A* and *CYP2A6*4D* alleles were expected to be amplified, respectively. The PCR product was digested with *Eco*81I restriction enzyme. The digestion patterns were determined by electrophoresis in a 2% agarose gel. The schematic RFLP patterns are shown in Fig. 4B.

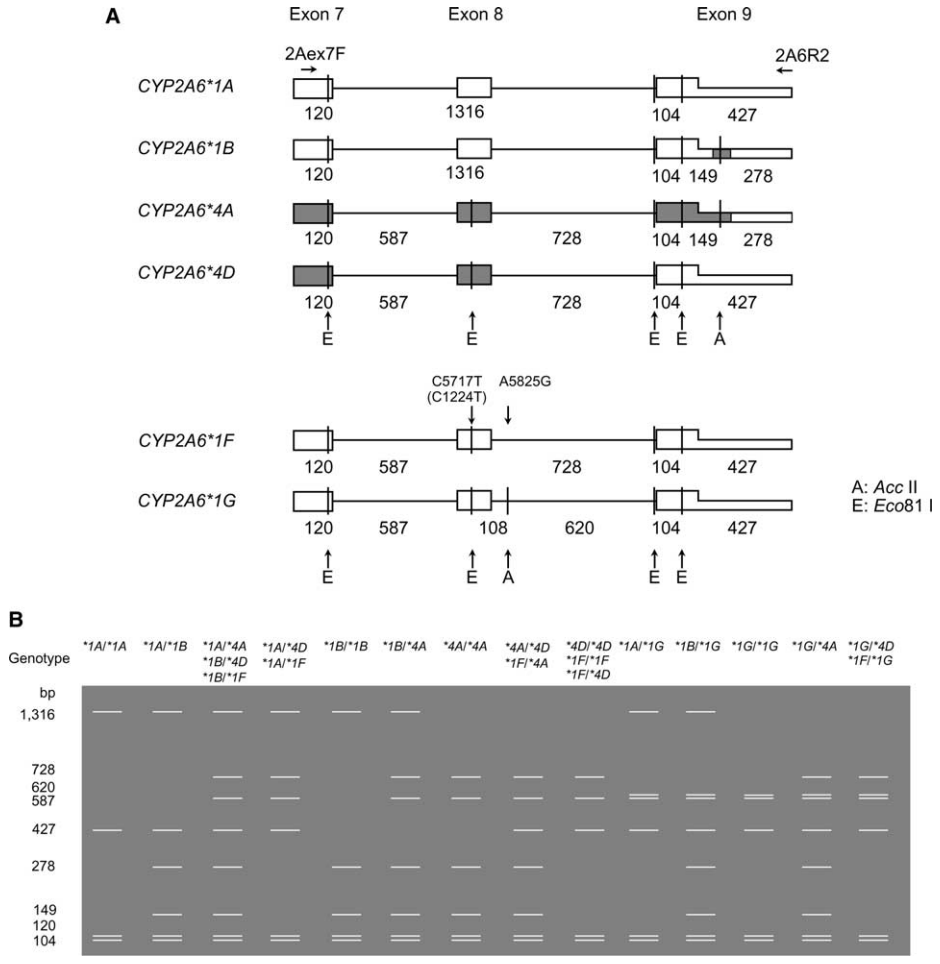

Fig. 2. Genotyping of *CYP2A6*1A, CYP2A6*1B, CYP2A6*4A, CYP2A6*4D, CYP2A6*1F*, and *CYP2A6*1G* alleles by PCR-RFLP. (A) Schematic structures of *CYP2A7* and *CYP2A6* genes. Dotted boxes and open boxes represent exons of *CYP2A7* and *CYP2A6*, respectively. Lines represent introns of each gene. PCR amplification was performed with the primer pairs indicated by horizontal arrows. The amplified DNA was double-digested by *Eco*81I and *Acc*II. The restriction sites of *Eco*81I and *Acc*II are indicated by vertical arrows of E and A, respectively. (B) Schematic PCR-RFLP patterns for different *CYP2A6* alleles. *CYP2A6*1A* yields 1,316, 427, 120, and 104 bp fragments, *CYP2A6*1B* yields 1,316, 278, 149, 120, and 104 bp fragments, *CYP2A6*4A* yields 728, 587, 278, 149, 120, and 104 bp fragments, *CYP2A6*4D* and *CYP2A6*1F* yield 728, 587, 427, 120, and 104 bp, and *CYP2A6*1G* yields 620, 587, 427, 120, 108, and 104 bp fragments.

2.6. Genotyping of the other CYP2A6 alleles

The genotypings of *CYP2A6*2*, *CYP2A6*3*, *CYP2A6*5*, *CYP2A6*6*, *CYP2A6*7*, *CYP2A6*8*, *CYP2A6*9*, *CYP2A6*10*, *CYP2A6*11*, and *CYP2A6*1X2* were also performed as described previously [11,25]. Genotyping of *CYP2A6*12* was carried out according to the method by Oscarson et al. [19] with slight modifications. Briefly, 2A6ex1 [19] or 2A7ex1 [19] and 2A6ex3R1 [19] primers were used. Genomic DNA samples (0.5 μg) were added to the PCR mixtures (25 μl) consisting of 1× PCR buffer, 1.5 mM $MgCl_2$, 0.4 μM of each primer, 250 μM dNTPs, and 1 U of *Taq* DNA polymerase. After an initial denaturation at 95 °C for 3 min, the amplification was performed by denaturation at 95 °C for 30 s, annealing at 63 °C for 30 s, and extension at 72 °C for 1.5 min for 28 cycles, followed by a final extension at 72 °C for 7 min. An aliquot (15 μl) of the PCR product was analyzed by electrophoresis with 0.8% agarose gel. The *CYP2A6*1* allele was amplified with the primer set of 2A6ex1 and 2A6ex3R1 (1,683 bp) and the *CYP2A6*12* allele was amplified with the primer set of 2A7ex1 and 2A6ex3R1 (1,598 bp).

3. Results

We applied a genotyping method of PCR-RFLP using primers of 2Aex7F and 2A6R2 to distinguish between the *CYP2A6*4A* and *CYP2A6*4D* alleles. As shown in Fig. 2A, *CYP2A6*1A* yields 1,316, 427, 120, and 104 bp fragments; *CYP2A6*1B* yields 1,316, 278, 149, 120, and 104 bp fragments; *CYP2A6*4A* yields 728, 587, 278, 149, 120, and 104 bp fragments; *CYP2A6*4D* yields 728, 587, 427, 120, and 104 bp fragments. The schematic RFLP patterns are shown in Fig. 2B. To confirm the identity of the *CYP2A6*4D* allele, DNA sequencing analysis was performed. The PCR product with the primers of 2Aex7F and 2A6R2 from a subject who was genotyped as *CYP2A6*1A*/*CYP2A6*4D* was subcloned into pT7Blue T-vector. As the results of the sequencing analyses, the allele that had been genotyped as *CYP2A6*4D* showed the same nucleotide sequence of *CYP2A6*1A* except for a SNP of C5717T in exon 8 (Accession No. AC008537, the A in the initiation codon is nucleotide +1). Therefore, the allele was not the *CYP2A6*4D*. The SNP of C5717T corresponds to C1224T on the cDNA sequence and is a synonymous mutation. The novel allele was termed *CYP2A6*1F*. The SNP of C5717T produces a recognition site of the restriction enzyme *Eco*81I. Therefore, the RFLP pattern of the *CYP2A6*1F* allele was the same as that of *CYP2A6*4D* allele.

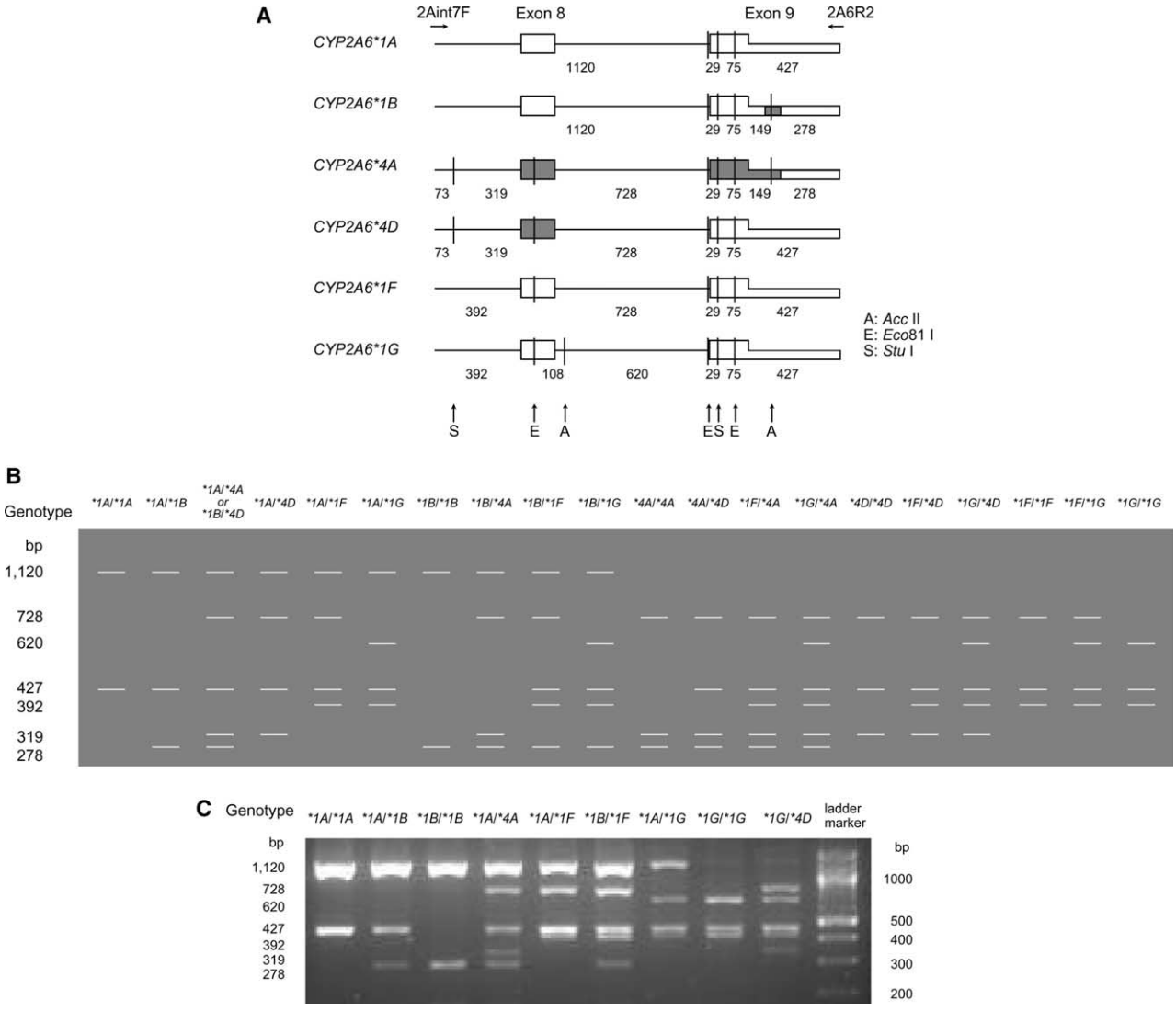

Fig. 3. Improved genotyping of *CYP2A6*1A*, *CYP2A6*1B*, *CYP2A6*4A*, *CYP2A6*4D*, *CYP2A6*1F*, and *CYP2A6*1G* alleles by PCR-RFLP. (A) Schematic structures of *CYP2A7* and *CYP2A6* genes. Dotted boxes and open boxes represent exons of *CYP2A7* and *CYP2A6*, respectively. Lines represent introns of each gene. Polymerase chain reaction amplification was performed with the primer pairs indicated by horizontal arrows. The amplified DNA was triple-digested by *Eco*81I, *Acc*II, and *Stu*I. The restriction sites are indicated by vertical arrows of E, A, and S, respectively. (B) Schematic PCR-RFLP patterns for different *CYP2A6* alleles. *CYP2A6*1A* yields 1,120, 427, 75, and 29 bp fragments, *CYP2A6*1B* yields 1,120, 278, 149, 75, and 29 bp fragments, *CYP2A6*4A* yields 728, 319, 278, 149, 75, 73, and 29 bp fragments, *CYP2A6*4D* yields 728, 427, 319, 75, 73, and 29 bp, *CYP2A6*1F* yields 728, 427, 392, 75, and 29 bp, and *CYP2A6*1G* yields 620, 427, 392, 108, 75, and 29 bp fragments. (C) Representative photograph of PCR-RFLP patterns for different *CYP2A6* alleles.

Accordingly, the *CYP2A6*1F* allele could be mistyped as the *CYP2A6*4D* allele.

In the process of genotyping with the RCR-RFLP method (Fig. 2), an unexpected 620 bp fragment was observed in several samples. To confirm the structure of the allele, the PCR product in which 620 bp fragments was observed was also subcloned into pT7Blue T-vector, and DNA sequencing analysis was performed. The allele has two SNPs of C5717T in exon 8 and A5825G in intron 8 creating the recognition site of *Eco*81I and *Acc*II, respectively. The allele was termed *CYP2A6*1G*.

To distinguish between the *CYP2A6*4D* and *CYP2A6*1F* alleles, the PCR-RFLP analysis was improved using the primers of 2Aint7F and 2A6R2 and triple-digestion with *Eco*81I, *Acc*II, and *Stu* I restriction enzymes. As shown in Fig. 3B, all genotypes demonstrate different RFLP patterns except for *CYP2A6*1A*/*CYP2A6*4A* and *CYP2A6*1B*/*CYP2A6*4D*. These two genotypes could be distinguished by second-step PCR as shown in Fig. 4. The sequences of the *CYP2A6*4D* allele genotyped with the improved PCR-RFLP method were confirmed by DNA sequencing analyses.

Genotyping of all *CYP2A6* alleles in 165 Caucasian and 94 African-American subjects was performed. In these subjects, *CYP2A6*1X2*, *CYP2A6*3*, *CYP2A6*5*, *CYP2A6*6*, *CYP2A6*7*, *CYP2A6*8*, *CYP2A6*10*, *CYP2A6*11*, *CYP2A6*12* were not observed. The allele frequencies of *CYP2A6* gene in Caucasians and African-Americans are summarized in Table 1. The allele frequencies of *CYP2A6*4A* were 3.0% and 0.5% in Caucasians and African-Americans, respectively. The allele frequencies of *CYP2A6*4D* were 0% and 0.5% in Caucasians and African-Americans, respectively. The *CYP2A6*1F* allele was found in 1.8% of Caucasians. There was a large ethnic difference in the

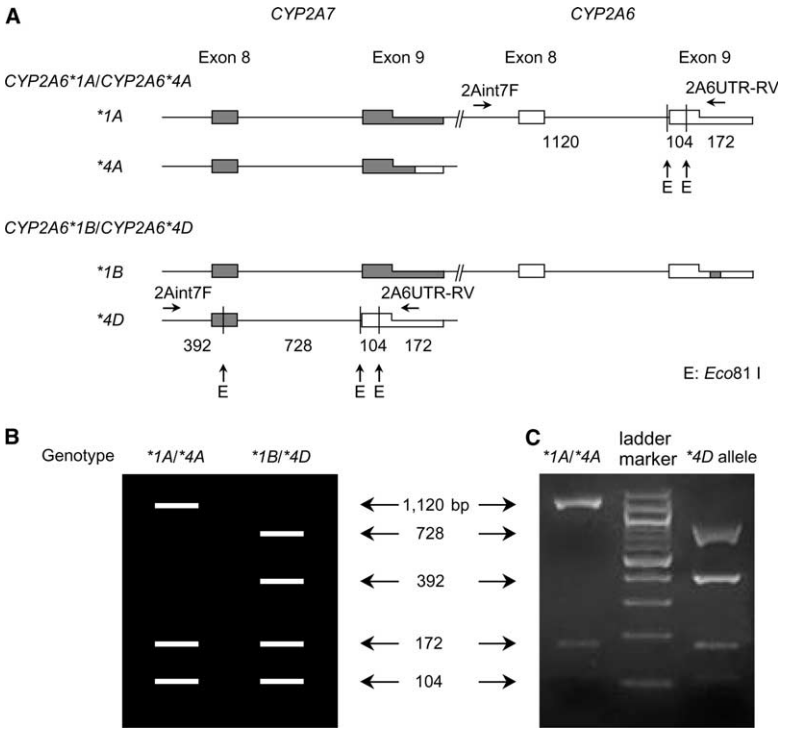

Fig. 4. Genotyping of *CYP2A6*1A*/*CYP2A6*4A* or *CYP2A6*1B*/*CYP2A6*4D* by PCR-RFLP. (A) Schematic structures of *CYP2A7* and *CYP2A6* genes. Dotted boxes and open boxes represent exons of *CYP2A7* and *CYP2A6*, respectively. Lines represent introns of each gene. PCR amplification was performed with the primer pairs indicated by horizontal arrows. The PCR products were digested by *Eco*81 I. The restriction sites of *Eco*81I are indicated by vertical arrows. (B) Schematic PCR-RFLP patterns for *CYP2A6*1A*/*CYP2A6*4A* or *CYP2A6*1B*/*CYP2A6*4D*. *CYP2A6*1A* yields 1,120, 172, and 104 bp fragments, *CYP2A6*4D* yields 728, 392, 172, and 104 bp fragments. (C) Representative photograph of PCR-RFLP patterns for *CYP2A6*1A*/*CYP2A6*4A* and *CYP2A6*4D* allele. In the subjects investigated in our study, there was no subject with the *CYP2A6*1B*/*CYP2A6*4D* allele. Therefore, PCR product for the *CYP2A6*4D* allele was prepared as follows to confirm the RFLP pattern: the PCR product with the genomic DNA genotyped *CYP2A6*1G*/*CYP2A6*4D* using 2Aint7F and 2A6UTR-RV primers was digested with *Acc*II. The PCR product derived from the *CYP2A6*1G* allele was digested to 896 and 500 bp fragments. Non-digested PCR product derived from the *CYP2A6*4D* allele was prepared from the agarose gel.

Table 1
CYP2A6 genotypes and allele frequencies (%) in 165 Caucasians and 94 African-Americans

Genotype	Number of subjects	
	Caucasians ($n = 165$)	African-Americans ($n = 94$)
*CYP2A6*1A*/*CYP2A6*1A*	55 (33.3)	39 (41.5)
*CYP2A6*1A*/*CYP2A6*1B*	49 (29.7)	18 (19.1)
*CYP2A6*1A*/*CYP2A6*1F*	3 (1.8)	0 (0)
*CYP2A6*1A*/*CYP2A6*1G*	0 (0)	17 (18.1)
*CYP2A6*1B*/*CYP2A6*1B*	15 (9.1)	0 (0)
*CYP2A6*1B*/*CYP2A6*1F*	3 (1.8)	0 (0)
*CYP2A6*1B*/*CYP2A6*1G*	0 (0)	1 (1.1)
*CYP2A6*1G*/*CYP2A6*1G*	2 (1.2)	2 (2.1)
*CYP2A6*1A*/*CYP2A6*2*	2 (1.2)	0 (0)
*CYP2A6*1A*/*CYP2A6*4A*	9 (5.5)	1 (1.1)
*CYP2A6*1B*/*CYP2A6*4A*	1 (0.6)	0 (0)
*CYP2A6*1G*/*CYP2A6*4D*	0 (0)	1 (1.1)
*CYP2A6*1A*/*CYP2A6*9*	16 (9.7)	11 (11.7)
*CYP2A6*1B*/*CYP2A6*9*	8 (4.8)	2 (2.1)
*CYP2A6*1G*/*CYP2A6*9*	0 (0)	2 (2.1)
*CYP2A6*2*/*CYP2A6*9*	2 (1.2)	0 (0)
Allele	Number of alleles	
	Caucasians ($n = 330$)	African-Americans ($n = 188$)
*CYP2A6*1A*	189 (57.3)	125 (66.5)
*CYP2A6*1B*	91 (27.6)	21 (11.2)
*CYP2A6*1F*	6 (1.8)	0 (0)
*CYP2A6*1G*	4 (1.2)	25 (13.3)
*CYP2A6*2*	4 (1.2)	0 (0)
*CYP2A6*4A*	10 (3.0)	1 (0.5)
*CYP2A6*4D*	0 (0)	1 (0.5)
*CYP2A6*9*	26 (7.9)	15 (8.0)

allele frequencies of the *CYP2A6*1G* in Caucasians (1.2%) and African-Americans (13.3%). The allele frequencies of *CYP2A6*2* and *CYP2A6*9* were consistent with previous reports [10,15,18].

4. Discussion

Both *CYP2A6*4A* and *CYP2A6*4D* are deleted alleles of the *CYP2A6* gene. An unequal crossover junction is located in the 3′-flanking region in the *CYP2A6*4A* allele, whereas the junction is located in either intron 8 or exon 9 in the *CYP2A6*4D* allele [15]. With the previous two-step PCR method using 2A7ex8F and 2A6R2 primers [15], it was impossible to distinguish the *CYP2A6*4D* allele from the *CYP2A6*4A* allele. In the present study, we developed a PCR-RFLP method for distinguishing between the *CYP2A6*4A* and *CYP2A6*4D* alleles. In the process of genotyping, two novel alleles of *CYP2A6*1F* and *CYP2A6*1G* were found. These alleles have a synonymous SNP of C5717T in exon 8. It is considered that the enzymatic activities of these alleles might be similar to those of the wild-type, which should be confirmed in the near future. Since the *CYP2A6*1F* allele demonstrated the same RFLP pattern of *CYP2A6*4D*, it would cause a mistyping. The improved genotyping method developed in the present study makes it possible to distinguish between the *CYP2A6*4A*, *CYP2A6*4D* and *CYP2A6*1F* alleles.

In the results of the genotyping of 165 Caucasians, the allele frequencies of *CYP2A6*4A* and *CYP2A6*4D* were 3.0% and 0%, respectively. In 94 African-Americans, the allele frequencies of *CYP2A6*4A* and *CYP2A6*4D* were both 0.5%. This is the first study to report separately the allele frequencies of the *CYP2A6*4A* and *CYP2A6*4D* alleles. Furthermore, we confirmed that the subjects genotyped with *CYP2A6*4* in Japanese and Koreans in our previous studies [22,23] were all *CYP2A6*4A* using the novel method established in the present study. Thus, the *CYP2A6*4D* allele was not found and the allele frequencies of *CYP2A6*4A* in Japanese and Koreans were 20.1% and 11.0%, respectively [22,23]. It has been reported that the allele frequency of *CYP2A6*4* is 1.0% in Finns [15], 0.5% in Spaniards [15], 15.1% in Chinese [15], and 1.18% in Caucasians [18]. With the PCR-RFLP method established in the present study, it will be possible to know the exact allele frequencies of *CYP2A6*4A* and *CYP2A6*4D* in these populations.

In our previous studies, the effects of genetic polymorphisms of the human *CYP2A6* gene on in vivo nicotine metabolism were confirmed [11,21–25,27]. Furthermore, a relationship between the *CYP2A6* genotype and smoking habits [18,28] as well as the incidence of lung cancer [29] has been indicated. Since CYP2A6 can metabolize pharmaceutical agents, the genetic polymorphism of the *CYP2A6* gene would be clinically important in so far as certain drugs are specifically metabolized by CYP2A6.

In conclusion, we developed a new genotyping method for distinguishing between two different whole deleted alleles, *CYP2A6*4A* and *CYP2A6*4D* as well as novel *CYP2A6*1F* and *CYP2A6*1G* alleles. Since the enzymatic activities of *CYP2A6*4D* and *CYP2A6*1F* are considered to be different, the mis-genotyping would show an apparent inconsistency between the genotype and phenotype.

Acknowledgements: The enthusiasm and research support of Tracy Jones, RN, Phyllis Klein, RN, and Ladonna Gaines, Washington University Center for Clinical Studies, and Margaret Ameyaw, MD was greatly appreciated. This study was supported by a grant from Japan Health Sciences Foundation with Research on Health Science Focusing on Drug Innovation, by an SRF Grant for Biomedical Research in Japan, and by Philip Morris Incorporated. We acknowledge Mr. Brent Bell for reviewing the manuscript.

References

[1] Oscarson, M. (2001) Genetic polymorphisms in the cytochrome P450 2A6 (*CYP2A6*) gene: implications for interindividual differences in nicotine metabolism. Drug Metab. Dispos. 29, 91–95.
[2] Nakajima, M., Yamamoto, T., Nunoya, K.-I., Yokoi, T., Nagashima, K., Inoue, K., Funae, Y., Shimada, N., Kamataki, T. and Kuroiwa, Y. (1996) Role of human cytochrome P4502A6 in *C*-oxidation of nicotine. Drug Metab. Dispos. 24, 1212–1217.
[3] Nakajima, M., Yamamoto, T., Nunoya, K., Yokoi, T., Nagashima, K., Inoue, K., Funae, Y., Shimada, N., Kamataki, T. and Kuroiwa, Y. (1996) Characterization of CYP2A6 involved in 3′-hydroxylation of cotinine in human liver microsomes. J. Pharmacol. Exp. Ther. 277, 1010–1015.
[4] Oscarson, M., McLellan, R.A., Gullstén, H., Agúndez, J.A.G., Benítez, J., Rautio, A., Raunio, H., Pelkonen, O. and Ingelman-Sundberg, M. (1999) Identification and characterization of novel polymorphisms in the *CYP2A* locus: implications for nicotine metabolism. FEBS Lett. 460, 321–327.
[5] Fernandez-Salguero, P., Hoffman, S.M.G., Cholerton, S., Mohrenweiser, H., Raunio, H., Rautio, A., Pelkonen, O., Huang, J.-D., Evans, W.E. and Idle, J.R. (1995) A genetic polymorphism in coumarin 7-hydroxylation: sequence of the human *CYP2A* genes and identification of variant *CYP2A6* alleles. Am. J. Hum. Genet. 57, 651–660.
[6] Yamano, S., Tatsuno, J. and Gonzalez, F.J. (1990) The CYP2A3 gene product catalyzes coumarin 7-hydroxylation in human liver microsomes. Biochemistry 29, 1322–1329.
[7] Kitagawa, K., Kunugita, N., Kitagawa, M. and Kawamoto, T. (2001) *CYP2A6*6*, a novel polymorphism in cytochrome P450 2A6, has a single amino acid substitution (R128Q) that inactivates enzymatic activity. J. Biol. Chem. 276, 17830–17835.
[8] Ariyoshi, N., Sawamura, Y. and Kamataki, T. (2001) A novel single nucleotide polymorphism altering stability and activity of CYP2A6. Biochem. Biophys. Res. Commun. 281, 810–814.
[9] Daigo, S., Takahashi, Y., Fujieda, M., Ariyoshi, N., Yamazaki, H., Koizumi, W., Tanabe, S., Saigenji, K., Nagayama, S., Ikeda, K., Nishioka, Y. and Kamataki, T. (2002) A novel mutant allele of the *CYP2A6* gene (*CYP2A6*11*) found in a cancer patient who showed poor metabolic phenotype towards tegafur. Pharmacogenetics 12, 299–306.
[10] Pitarque, M., von Richter, O., Oke, B., Berkkan, H., Oscarson, M. and Ingelman-Sundberg, M. (2001) Identification of a single nucleotide polymorphism in the TATA box of the *CYP2A6* gene: impairment of its promoter activity. Biochem. Biophys. Res. Commun. 284, 455–460.
[11] Yoshida, R., Nakajima, M., Watanabe, Y., Kwon, J.-T. and Yokoi, T. (2002) Genetic polymorphisms in human *CYP2A6* gene causing impaired nicotine metabolism. Br. J. Clin. Pharmacol. 54, 511–517.
[12] Xu, C., Rao, Y.S., Hoffmann, E., Jones, J., Sellers, E.M. and Tyndale, R.F. (2002) An in vivo pilot study characterizing the new *CYP2A6*7*, **8*, and **10* alleles. Biochem. Biophys. Res. Commun. 290, 318–324.
[13] Nunoya, K., Yokoi, T., Kimura, K., Kainuma, T., Satoh, K., Kinoshita, M. and Kamataki, T. (1999) A new *CYP2A6* gene deletion responsible for the in vivo polymorphic metabolism of (+)-*cis*-3,5-dimethyl-2-(3-pyridyl)thiazolidin-4-one hydrochloride in humans. J. Pharmacol. Exp. Ther. 289, 437–442.
[14] Nunoya, K., Yokoi, T., Takahashi, Y., Kimura, K., Kinoshita, M. and Kamataki, T. (1999) Homologous unequal cross-over within the human *CYP2A* gene cluster as a mechanism for the deletion of entire *CYP2A6* gene associated with the poor metabolizer phenotype. J. Biochem. 126, 402–407.

[15] 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) Characterization and PCR-based detection of a *CYP2A6* gene deletion found at a high frequency in a Chinese population. FEBS Lett. 448, 105–110.
[16] Ariyoshi, N., Takahashi, Y., Miyamoto, M., Umetsu, Y., Daigo, S., Tateishi, T., Kobayashi, S., Mizorogi, Y., Loriot, M.-A., Stücker, I., Beaune, P., Kinoshita, M. and Kamataki, T. (2000) Structural characterization of a new variant of the*CYP2A6* gene (*CYP2A6*1B*) apparently diagnosed as heterozygotes of *CYP-2A6*1A* and *CYP2A6*4C*. Pharmacogenetics 10, 687–693.
[17] Ariyoshi, N., Sekine, H., Saito, K. and Kamataki, T. (2002) Characterization of a genotype previously designated as *CYP2A6* D-type: *CYP2A6*4B*, another entire gene deletion allele of the CYP2A6 gene in Japanese. Pharmacogenetics 12, 501–504.
[18] Rao, Y., Hoffmann, E., Zia, M., Bodin, L., Zeman, M., Sellers, E.M. and Tyndale, R.F. (2000) Duplications and defects in the *CYP2A6* gene: identification, genotyping, and in vivo effects on smoking. Mol. Pharmacol. 58, 747–755.
[19] Oscarson, M., McLellan, R.A., Asp, V., Ledesma, M., Ruiz, M.L.B., Sinues, B., Rautio, A. and Ingelman-Sundberg, M. (2002) Characterization of a novel *CYP2A7*/*CYP2A6* hybrid allele (*CYP2A6*12*) that causes reduced CYP2A6 activity. Hum. Mutat. 20, 275–283.
[20] Kiyotani, K., Fujieda, M., Yamazaki, H., Shimada, T., Guengerich, F.P., Parkinson, A., Nakagawa, K., Ishizaki, T. and Kamataki, T. (2002) Twenty one novel single nucleotide polymorphisms (SNPs) of the CYP2A6 gene in Japanese and Caucasians. Drug Metab. Pharmacokin. 17, 482–487.
[21] Nakajima, M., Yamagishi, S., Yamamoto, H., Yamamoto, T., Kuroiwa, Y. and Yokoi, T. (2000) Deficient cotinine formation from nicotine is attributed to the whole deletion of the *CYP2A6* gene in humans. Clin. Pharmacol. Ther. 67, 57–69.
[22] Nakajima, M., Kwon, J., Tanaka, N., Zenta, T., Yamamoto, Y., Yamamoto, H., Yamazaki, H., Yamamoto, T., Kuroiwa, Y. and Yokoi, T. (2001) Relationship between interindividual differences in nicotine metabolism and *CYP2A6* genetic polymorphism in humans. Clin. Pharmacol. Ther. 69, 72–78.
[23] Kwon, J.-T., Nakajima, M., Chai, S., Yom, Y.-K., Kim, H.-K., Yamazaki, H., Sohn, D.-R., Yamamoto, T., Kuroiwa, Y. and Yokoi, T. (2001) Nicotine metabolism and *CYP2A6* allele frequencies in Koreans. Pharmacogenetics 11, 317–323.
[24] Yoshida, R., Nakajima, M., Watanabe, Y., Kwon, J.-T. and Yokoi, T. (2002) Genetic polymorphisms in human *CYP2A6* gene causing impaired nicotine metabolism. Br. J. Clin. Pharmacol. 54, 511–517.
[25] Yoshida, R., Nakajima, M., Nishimura, K., Tokudome, S., Kwon, J.-T. and Yokoi, T. (2003) Effects of polymorphism in promoter region of human *CYP2A6* gene (*CYP2A6*9*) on expression level of mRNA and enzymatic activity in vivo and in vitro. Clin. Pharmacol. Ther. 74, 69–96.
[26] Benowitz, N.L, Griffin, C. and Tyndale, R. (2001) Deficient C-oxidation of nicotine continued. Clin. Pharmacol. Ther. 70, 567.
[27] Nakajima, M., Kuroiwa, Y. and Yokoi, T. (2002) Interindividual differences in nicotine metabolism and genetic polymorphisms of human CYP2A6. Drug Metab. Rev. 34, 865–877.
[28] Pianezza, M.L., Sellers, E.M. and Tyndale, R.F. (1998) Nicotine metabolism defect reduces smoking. Nature 393, 750.
[29] Miyamoto, M., Umetsu, Y., Dosaka-Akita, H., Sawamura, Y., Yokota, J., Kunitoh, H., Nemoto, N., Sato, K., Ariyoshi, N. and Kamataki, T. (1999) *CYP2A6* gene deletion reduces susceptibility to lung cancer. Biochem. Biophys. Res. Commun. 261, 658–660.