

**Title page**

A null allele impairs function of *CYP2C76* gene in cynomolgus monkeys: a possible genetic tool for generation of a better animal model in drug metabolism.

Yasuhiro Uno, Hiroko Sakuraba, Shotaro Uehara, Takayuki Kumano, Kiyomi Matsuno, Chika Nakamura, Go Kito, Tetsuya Kamataki, and Ryoichi Nagata

Laboratories of Translational Research (Y.U., H.S., T.Ku., G.K., R.N.) and Drug Metabolism (T.Ka.), Graduate School of Pharmaceutical Sciences, Hokkaido University, Hokkaido, Japan; Shin Nippon Biomedical Laboratories, Ltd., Pharmacokinetics and Bioanalysis Center, Wakayama, Japan (Y.U., S.U., K.M., C.N.); Shin Nippon Biomedical Laboratories, Ltd., Drug Safety Research Laboratories, Kagoshima, Japan (G.K., R.N.).

**Running title page**

**Running title:** *CYP2C76* null allele in cynomolgus monkeys

**Corresponding Author:** Dr. Yasuhiro Uno

Shin Nippon Biomedical Laboratories (SNBL), Ltd., Pharmacokinetics and Bioanalysis Center (PBC),

Genome Research Group, 16-1 Minami Akasaka, Kainan 642-0017, Japan.

Phone: +81-73-483-8881, Fax: +81-73-483-7377, *E-mail address:* [uno-yasuhiro@snbl.co.jp](mailto:uno-yasuhiro@snbl.co.jp)

<b>Number of Text Pages :</b>	13
<b>Number of Tables :</b>	2
<b>Number of Figures :</b>	1
<b>Number of References :</b>	16
<b>Number of Words in Abstract :</b>	202
<b>Number of Words in Introduction :</b>	450
<b>Number of Words in Results and Discussion :</b>	914

**Abbreviations:** CYP, cytochrome P450; EST, expressed sequence tag; PCR, polymerase chain reaction; SRS, substrate recognition site.

## Abstract

The monkey *CYP2C76* gene does not correspond to any of the human *CYP2C* genes and its enzyme is at least partly responsible for the species difference occasionally seen in drug metabolism between monkeys and humans. To establish a line and/or lines of monkeys that are expected to show metabolic patterns highly similar to humans, we set out to find monkeys lacking *CYP2C76* activity. By genetic screening of 73 monkeys and a database search of expressed sequence tags, we found a total of 10 non-synonymous genetic variants in the coding region of *CYP2C76* including a null genotype (c.449TG>A). Some of the variants were differently distributed between two animal groups originating from different geographical regions (Indochina and Indonesia). After screening 170 additional genomic samples, we identified a total of 8 animals (6 males and 2 females) heterozygous for c.449TG>A, which could be used for establishing a homozygous line. If the homozygotes show drug-metabolizing properties more similar to humans than wild-type monkeys, the homozygotes may serve as a better animal model for drug metabolism. The data presented in this paper provide the essential genetic information of a successful study using cynomolgus monkeys and present a possible tool for generation of better drug metabolism animal models.

## Introduction

Monkeys, especially macaques which include cynomolgus monkeys (*Macaca fascicularis*), are frequently used to predict the metabolic fate of drugs in humans due to their similar pharmacokinetics to humans. However, differences are occasionally seen between the two species for some drugs (Stevens et al., 1993; Sharer et al., 1995; Weaver et al., 1999). To predict the metabolism of drugs and the influence of drug-drug interactions in humans, species differences in drug-metabolizing enzymes including cytochrome P450s (CYPs) should be clarified and taken into account when extrapolating monkey data to humans.

CYP is one of the most important drug-metabolizing enzymes and it is known to form a superfamily consisting of a large number of subfamilies (Nelson et al., 2004). One of the subfamilies, the CYP2C subfamily including CYP2C8, CYP2C9, and CYP2C19 in humans is important for drug metabolism because the CYP2Cs metabolize approximately 20% of all prescribed drugs, such as tolbutamide, phenytoin, warfarin, and ibuprofen (Goldstein, 2001). For macaques, other groups as well as ours have identified cDNAs for CYP2C20, CYP2C43, CYP2C75, and CYP2C76 from cynomolgus or rhesus monkeys (Komori et al., 1992; Matsunaga et al., 2002; Mitsuda et al., 2006; Uno et al., 2006). Among these, the failure to identify a human gene homologous to CYP2C76 at the corresponding region of the human genome as well as the absence of protein in the human liver strongly suggest that none of the human genes are orthologous to *CYP2C76* (Uno et al., 2006). *CYP2C76* was, at least partly, responsible for the species difference in the metabolism of pitavastatin between monkeys and humans (Uno et al., 2007a). Moreover, investigation of other 11 CYPs isolated from a liver cDNA library revealed that these CYPs were, unlike *CYP2C76*, not apparent species-specific, raising the possibility that *CYP2C76* might be one of a few species-specific CYPs substantially expressed in cynomolgus monkey liver (Uno et al., 2007b). These results suggest the importance of *CYP2C76* for species difference in drug metabolism between monkeys and humans.

One way to improve accuracy in extrapolating animal data to humans is to establish better animal models. Eliminating the function of the species-specific genes such as *CYP2C76* in monkeys might

lead to drug-metabolizing properties more similar to humans. Inactivation of the gene itself can be accomplished by gene targeting, but the technique has not been successfully applied to non-human primate species (Norgren, 2004). As in humans, monkeys have a diverse genetic background with a variety of genetic polymorphisms. Functionally null or defective alleles have been detected in the human *CYP* genes such as *CYP1B1*, *CYP2A6*, *CYP2B6*, *CYP2C8*, *CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP3A4*, and *CYP3A5* (see <http://www.imm.ki.se/CYPalleles/>). In this study, therefore, we paid efforts to comprehensively identify genetic variants to identify null and/or defective alleles present in *CYP2C76*.

## Methods

### Animals and genomic DNA extraction

Genomic DNA was prepared from whole blood samples using the PUREGENE<sup>®</sup> DNA isolation kit (Gentra Systems, Minneapolis, MN) according to the manufacturer's instructions. The blood samples used in this study were collected from 243 cynomolgus monkeys (39 from Indochina, 204 from Indonesia). The study was approved by the local ethics committee.

### Identification and analysis of *CYP2C76* genetic variants

To identify genetic variants and to determine allele frequencies of the identified variants, we used the genome samples of 73 cynomolgus monkeys (39 from Indochina, 34 from Indonesia). For further genotype screening, 170 additional genomes from animals of Indonesian origin were analyzed. Each exon as well as flanking DNA was amplified by PCR, followed by direct sequencing of the PCR products. PCR was carried out in 20- $\mu$ L reactions containing 1 ng of genomic DNA, 5 pmole of forward and reverse primers, and 1 unit of AmpliTaq<sup>™</sup> Gold DNA polymerase (AppliedBiosystems, Foster City, CA). The amplification was performed on the MJ Research thermal cycler (BioRad, Hercules, CA) with an initial denaturation at 95°C for 10 min and 30 cycles of 20 s at 95°C, 30 s at 55°C, and 1 min at 72°C, followed by a final extension step of 10 min at 72°C. Sequencing was conducted with the primer used for PCR using ABI PRISM<sup>®</sup> BigDye<sup>™</sup> Terminator v3.0 Ready

Reaction Cycle Sequencing Kit (Applied Biosystems), followed by electrophoresis with the ABI PRISM<sup>®</sup> 3730 DNA Analyzer (Applied Biosystems). The primer pairs used for PCR and sequencing are listed in Table 1. Sequence data were analyzed using DNASIS Pro (Hitachi Software, Tokyo, Japan), which was also utilized for the identification of genetic variants from an in-house EST database.

### **Preparation of expression plasmids and protein expression**

For the functional characterization of genetic variants, the construction of expression plasmids, protein expression, and measurement of CYP proteins were performed as previously described (Uno et al., 2006). A mutation was introduced into the CYP2C76 expression plasmid using the QuikChange<sup>™</sup> XL II kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions. The primer pairs employed are listed in Table 1. The entire sequence of each insert was confirmed by sequencing.

### **Measurement of tolbutamide 4-hydroxylation**

To measure the drug-metabolizing activity of CYP2C76 variants, tolbutamide methyl-hydroxylation was measured to which CYP2C76 shows the drug-metabolizing activity (Uno et al., 2006; 2007a). To prepare the reaction mixture, tolbutamide was pre-incubated in 250 mM potassium phosphate buffer solution (pH 7.4) with the partially purified recombinant CYP proteins at 37°C for 5 min. The reaction was initiated by the addition of NADPH regenerating system (BD Biosciences, San Jose, CA) in a total volume of 0.5 ml. The incubations were carried out at 37°C in triplicate at linear conditions for 30 min using tolbutamide at 5 concentrations between 0.25 and 3.0 mM. Ice-cold 70% perchloric acid (50  $\mu$ l) was added to stop the reaction. An internal standard solution and ultrapure water were added to the reaction mixture and the resulting mixture was centrifuged (7500g, 4°C, 10 min). The metabolites in the supernatant were analyzed by LC/MS/MS (Shimadzu, Kyoto, Japan). The Michaelis-Menten parameters of the apparent  $K_m$  and  $V_{max}$  were determined using WinNonlin (version 4.0, Pharsight, Mountain View, CA).

## Results and Discussion

Our initial survey of 73 genome samples (39 from Indochina, 34 from Indonesia) and our in-house EST database identified a number of variants over all exons and their vicinities of *CYP2C76* in cynomolgus monkeys. To describe genetic variants, the *CYP2C76* cDNA sequence (DQ074807) originally identified (Uno et al., 2006) was regarded as a reference sequence in this study due to the unavailability of a consensus wild-type sequence for *CYP2C76*. Since we were interested in the null or defective alleles to generate an animal model, only non-synonymous variants were listed (Table 2). A total of 10 non-synonymous variants were identified, of which 3 variants, c.707A>T, c.883A>G, and c.1292A>G, were located in the regions important for protein function such as SRS (Gotoh et al., 1992) and heme-binding region. Because of the potential relevance of these variants, their influence on protein function was investigated using recombinant proteins heterologously expressed in *E. coli*. The drug-metabolizing capacity of the recombinant proteins using tolbutamide as a substrate indicated that these 3 variants showed an appreciable reduction in metabolic activity as compared to the *CYP2C76* protein of the reference sequence (Table 2).

Importantly, we found a null allele c.449TG>A in exon 3 (Table 2), causing a frameshift and a premature termination codon in the same exon. The sequencing chromatogram for this allele is described in Fig. 1A. With this allele, approximately three-fourths of the entire protein is predicted to be truncated including the heme-binding region and all the SRSs except for SRS1, indicating that the protein, even if translated, would not be functional. Indeed, the CYP-specific peaks with a maximum at 450 nm observed for the protein of reference sequence were not found for the protein with this genotype by reduced CO-difference spectra (Fig. 1B), strongly indicating the c.449TG>A to be a null allele. Because of the importance of c.449TG>A for establishing a better animal model, further screening was carried out using additional genomic samples of 170 cynomolgus monkeys from Indonesia, since none of the animals from Indochina analyzed possessed this allele (Table 2). Overall, 8 individuals were found heterozygous with this mutation including 6 males and 2 females; the allele frequency was estimated to be 0.020 for the animals of Indonesian origin analyzed in this

study.

The allele frequency of the other variants identified was determined with the genome samples of the 73 animals used for the discovery of the genetic variants. The genetic variants such as c.707A>T and c.928A>T were found to be major alleles in the animal groups analyzed, likely reflecting the genetic difference of the animals used to identify CYP2C76 cDNA and those employed in this study. Some of the variants including c.449TG>A, were distributed differently between the two animal groups. One reason for this distinction could be the regional difference in the allele frequency, since the two animal groups originated from different geographic regions, Indochina and Indonesia, and are separated by an ocean. Similarly, the regional difference of allele frequency has been revealed for the rhesus monkeys of the Chinese and Indian origin as well (Ferguson et al., 2007; Rhesus Macaque Genome Sequencing and Analysis Consortium, 2007). Alternatively, the difference might be simply attributable to the different colonies from which the animals analyzed in this study were derived. One allele, c.1292A>G, detected by searching the EST database was not found in the 73 animals analyzed in this study (Table 2). The reason for this discrepancy may be accounted for by assuming that the animals used in this study and those employed for EST sequencing were derived from different colonies. Further screening of animals from various colonies in different regions should give a clearer answer to this question. Considering the possibility that genetic variants such as c.449TG>A could substantially affect drug metabolism in monkeys, the origin of animals used is one of the most important factors for a successful study on drug metabolism.

Inactivation of species-specific genes such as *CYP2C76* could lead to a better animal model in monkeys. However, the techniques of gene knockout or knockdown *in vivo*, have not been available in this species (Norgren, 2004). Our results present an alternative way to produce animals lacking the function of a specific protein by utilizing genetic polymorphisms such as null or defective alleles in monkeys. It should be noted that in our case this is only possible if *CYP2C76* does not play crucial roles in embryonic or fetal development. This might be one of the reasons for the failure to identify homozygotes after screening 243 animals. For example, *CYP26A1* protects tissues from overexposure to all-*trans* retinoic acid by metabolizing (degrading) it, otherwise, embryos cannot



survive as evidenced by prenatal death of the CYP26A1 null mouse (Niederreither et al., 2002). CYP2C76 might play a similar developmental role if CYP2C76 is expressed in embryos or fetuses. Nevertheless, this approach is not only limited to *CYP2C76* but also applicable for many other genes, and should be useful for generation of better animal models for humans in a variety of physiological investigations.

In this study, our initial screening for genetic variants in cynomolgus *CYP2C76* has successfully identified a total of 8 non-synonymous variants. These include the null allele (c.449TG>A), which resulted in a frameshift and premature termination codon introduced in exon 3. So far, 6 males and 2 females have been found to be heterozygous for this allele. These animals can now be used for generating homozygotes through mating. If the homozygotes show drug-metabolizing properties more similar to humans, they may serve as a better animal model for drug metabolism.

## References

- Ferguson B, Street SL, Wright H, Pearson C, Jia Y, Thompson SL, Allibone P, Dubay CJ, Spindel E and Norgren RB Jr (2007) Single nucleotide polymorphisms (SNPs) distinguish Indian-origin and Chinese-origin rhesus macaques (*Macaca mulatta*). *BMC Genomics* **8**:43.
- Goldstein JA (2001) Clinical relevance of genetic polymorphisms in the human CYP2C subfamily. *Br J Clin Pharmacol* **52**:349-355.
- Gotoh O (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J Biol Chem* **267**:83-90.
- Komori M, Kikuchi O, Sakuma T, Funaki J, Kitada M and Kamataki T (1992) Molecular cloning of monkey liver cytochrome P-450 cDNAs: similarity of the primary sequences to human cytochromes P-450. *Biochim Biophys Acta* **1171**:141-146.
- Matsunaga T, Ohmori S, Ishida M, Sakamoto Y, Nakasa H and Kitada M (2002) Molecular cloning of monkey CYP2C43 cDNA and expression in yeast. *Drug Metab Pharmacokinet* **17**:117-124.
- Mitsuda M, Iwasaki M and Asahi S (2006) Cynomolgus monkey cytochrome P450 2C43: cDNA cloning, heterologous expression, purification and characterization. *J Biochem (Tokyo)* **139**:865-872.
- Nelson DR, Zeldin DC, Hoffman SM, Maltais LJ, Wain HM and Nebert DW (2004) Comparison of cytochrome P450 (*CYP*) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. *Pharmacogenetics* **14**:1-18.
- Niederreither K, Abu-Abed S, Schuhbaur B, Petkovich M, Chambon P and Dollé P (2002) Genetic evidence that oxidative derivatives of retinoic acid are not involved in retinoid signaling during mouse development. *Nat Genet* **31**:84-88.
- Norgren RB Jr (2004) Creation of non-human primate neurogenetic disease models by gene targeting and nuclear transfer. *Reprod Biol Endocrinol* **2**: 40.

- Rhesus Macaque Genome Sequencing and Analysis Consortium (2007) Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**:222-234.
- Sharer JE, Shipley LA, Vandenbranden MR, Binkley SN and Wrighton SA (1995) Comparisons of phase I and phase II *in vitro* hepatic enzyme activities of human, dog, rhesus monkey, and cynomolgus monkey. *Drug Metab Dispos* **23**:1231-1241.
- Stevens JC, Shipley LA, Cashman JR, Vandenbranden M and Wrighton SA (1993) Comparison of human and rhesus monkey *in vitro* phase I and phase II hepatic drug metabolism activities. *Drug Metab Dispos* **21**:753-760.
- Uno Y, Fujino H, Kito G, Kamataki T and Nagata R (2006) CYP2C76, a novel CYP in cynomolgus monkey, is a major CYP2C in liver, metabolizing tolbutamide and testosterone. *Mol Pharmacol* **70**:477-486.
- Uno Y, Kumano T, Kito G, Nagata R, Kamataki T and Fujino H (2007a) CYP2C76-mediated species difference in drug metabolism: A comparison of pitavastatin metabolism between monkeys and humans. *Xenobiotica* **37**:30-43.
- Uno Y, Hosaka S, Matsuno K, Nakamura C, Kito G, Kamataki T and Nagata R (2007b) Characterization of cynomolgus monkey cytochrome P450 (CYP) cDNAs: Is *CYP2C76* the only monkey-specific *CYP* gene responsible for species differences in drug metabolism? *Arch Biochem Biophys* **46**:98-105.
- Weaver RJ, Dickins M and Burke MD (1999) A comparison of basal and induced hepatic microsomal cytochrome P450 monooxygenase activities in the cynomolgus monkey (*Macaca fascicularis*) and man. *Xenobiotica* **29**:467-482.

## Footnotes

Request reprints to:

Shin Nippon Biomedical Laboratories (SNBL), Ltd., Pharmacokinetics and Bioanalysis Center (PBC),  
Genome Research Group, 16-1 Minami Akasaka, Kainan 642-0017, Japan.

Phone: +81-73-483-8881, Fax: +81-73-483-7377, *E-mail address:* [uno-yasuhiro@snbl.co.jp](mailto:uno-yasuhiro@snbl.co.jp)

The data presented here have been included in the dissertation of Yasuhiro Uno that was submitted to  
Graduate School of Veterinary Medicine of Hokkaido University in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy.

### Legends for figures

**Fig. 1.** Nucleotide sequence and protein expression of c.449TG>A for *CYP2C76* in cynomolgus monkeys. The nucleotides of reference sequence and mutant sequence for c.449TG>A are indicated as chromatograms (A). For c.449TG>A, sequence of cDNA clone isolated from heterozygous monkey liver is shown. CO-difference spectra did not show peaks at 450 nm for *CYP2C76* proteins of mutant sequences (B).

**Table 1**

The primers used for amplification and sequencing of exons and for directed mutagenesis.

<b>Exon/ variant</b>	<b>Sequence (5' → 3')*</b>	
<b>For amplification and sequencing of exons:</b>		
Exon 2	F	CCCTGTGATATAGTAGAGGCTGA
	R	CACTTATGATTTCTCAATTTCTTC
Exon 3	F	GAAAAGCCAGCCCTGAAA
	R	TGCCTGGCCTAAAGTATTTG
	S	TTCAGCCATCATCATATTTTGT
Exon 5	F	GTGATTGGCCTTGAAGTTGC
	R	GCCCCAGATAAGGAATCAG
Exon 6	F	CTGACTGGCTCTAGTGGGTTAG
	R	CCATCGCAGGGAGAATAATG
	S	TCTTCACCCATAGCACACTCTA
Exon 7	F	GAGGATTGAACGTCATGGTG
	R	GGCAGAGTGTGGAGCTATCA
Exon 9	F	CAATGTGCATGGACAGAAAAG
	R	TGCCTAGACAGGTAGATAGGAGTG
<b>For directed mutagenesis:</b>		
c.449TG>	F	CCGAATTCAAGAGGAAGCCTaTGTCTGGTGGAAAG
	R	CTTCCACCAGACAAGGCTTCCTCTTGAATTCGG
c.707A>T	F	GGAAGTCATAATGTGTTATTTAAAAATTTTGCTCTCCAAAGAAG
	R	CTTCTTTGGAGAGCAAAATTTTAAATAACACATTATGACTTCC
c.883A>G	F	CTTGGTTGCTACTATATGGGATATGTTTgGTGCAGGAACAGAG
	R	CTCTGTTCCCTGCACCAACATATCCCATATAGTAGCAACCAAG
c.1292A>G	F	GCTTTTTTCAGCAGGAAgAAGAATTTGTGCTGGAGAAGGC
	R	GCCTTCTCCAGCACAAATTCTTcTTCCTGCTGAAAAAGC

\* F, forward primer, R, reverse primer, S, sequence primer. Unless otherwise specified, either PCR primer was used for sequencing. In the primer sequence for directed mutagenesis, a nucleotide to be changed is indicated in lower case.

**Table 2**

The genetic variants identified in cynomolgus *CYP2C76*.

Variant	Exon	Nucleotide change <sup>a</sup>	Amino acid change	Location on cDNA	Allele frequency		Site	Kinetic parameters <sup>b</sup>	
					Indochina (n = 39)	Indonesia (n = 34)		K <sub>m</sub>	V <sub>max</sub>
c.178G>T	2	caaaa(G>T)attat	D60Y	178	0.513	0.500		N.D.	N.D.
c.193C>T	2	ctgtg(C>T)tcact	L65F	193	0.513	0.500		N.D.	N.D.
c.241G>A	2	gatac(G>A)aagca	E81K	241	0.513	0.500		N.D.	N.D.
c.449TG>A	3	agcct(TG>A)tgtct	Frameshift	449	0	0.020 <sup>d</sup>		N.D.	N.D.
c.707A>T	5	aaatt(A>T)tgctc	Y236F	707	0.936	1.000	SRS3	710	412
c.883A>G	6	tgttt(A>G)gtgca	S295G	883	0.013	0	SRS4	824	927
c.928A>T	6	gactg(A>T)tgctc	M310L	928	0.731	1.000		N.D.	N.D.
c.968G>A	7	agtgc(G>A)ggaag	R323Q	968	0.231	0		N.D.	N.D.
c.1292A>G <sup>c</sup>	9	agga(A>G)aagaa	K431R	1292	0	0	Heme-binding	836	1393
c.1436C>T	9	tatac(C>T)accct	P479L	1436	0.205	0.074		N.D.	N.D.

N.D., not determined.

<sup>a</sup> Nucleotide changes were detected by comparing with the *CYP2C76* cDNA sequence (GenBank accession no. DQ074807) as a reference sequence.

<sup>b</sup> K<sub>m</sub> (μM) and V<sub>max</sub> (pmol/min/nmol CYP) of each recombinant protein was determined with tolbutamide methyl-hydroxylation, for which the detection limit was 40 pmol/ml. The K<sub>m</sub> and V<sub>max</sub> for the *CYP2C76* of the reference sequence was 704 and 1570, respectively.

<sup>c</sup> This allele was identified by searching the EST database and was not found in any genome samples analyzed in this study.

<sup>d</sup> For c.449TG>A, 204 animals in Indonesian population were used to calculate allele frequency.

Fig. 1

