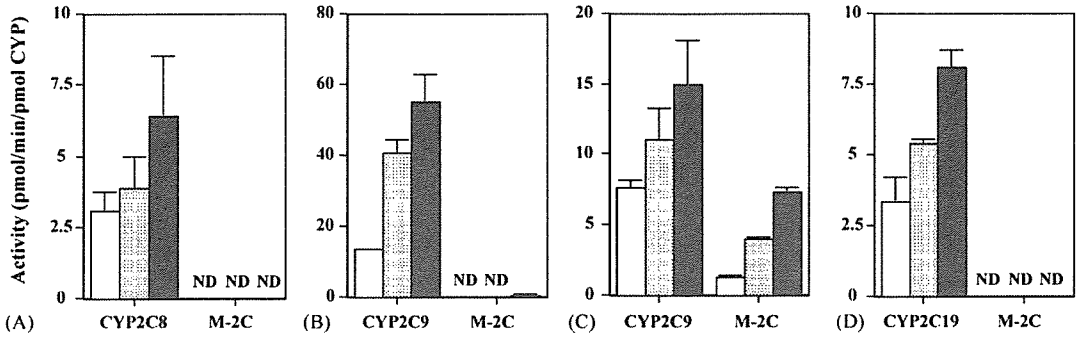

Fig. 4 – Comparison of various drug oxidation activities between P450 M-2C and human CYP2C enzymes. (A) PT (2.5, 5 and 10 μM) 6α-hydroxylation, (B) DF (5, 20 and 100 μM) 4′-hydroxylation, (C) TB (0.25, 1 and 2.5 mM) *p*-methylhydroxylation and (D) S-MP (10, 50 and 200 μM) 4′-hydroxylation. Open, dotted and hatched columns show the lowest, intermediate and highest concentrations, respectively. Each value represents the mean ± S.D. of three independent determinations. ND, not detectable.

key CYP2C20 (S53046), rhesus monkey CYP2C43 (AB212264), CYP2C74 (AY635462) and CYP2C75 (AY635463), respectively. The deduced amino acid sequence of P450 M-2C was highly identical to those of human CYP2C8 (87.1% identity), crab-eating monkey CYP2C20 (88.8%) and rhesus monkey CYP2C74 (89.0%).

3.2. *Expression of marmoset P450-M-2C protein in yeast cells*

The microsomal fraction was prepared from yeast cells expressing P450 M-2C, and the content of the recombinant holoenzyme was determined by reduced CO-difference spectroscopy (Fig. 3, left panel). The spectrum showed a Soret peak at 450 nm and a negligible level of peak at 420 nm. The content of P450 M-2C was calculated to be 133 pmol/mg protein (the mean value of two independent determinations). In Western blot analysis using polyclonal antibodies raised against human CYP2C19 (Fig. 3, right panel), microsomal fractions from yeast cells expressing P450 M-2C (lane no. 5) and from the marmoset liver (lane no. 6) exhibited a single protein band with a molecular weight similar to that of recombinant CYP2C19 (lane no. 4). In contrast, the pooled human liver microsomal fraction (lane no. 1) showed a major protein band whose molecular weight was similar to that of recombinant CYP2C9 (lane no. 3) and additional three faint protein bands, two of which exhibited similar mobilities to recombinant CYP2C8 (lane no. 2) and CYP2C19 (lane no. 4).

Table 2 – Kinetic parameters for TB *p*-methylhydroxylation by microsomal fractions from yeast cells expressing marmoset and human CYP enzymes and from human and marmoset livers

Enzyme source	K_m (μM)	V_{max}	V_{max}/K_m
Recombinant enzyme[a]			
P450 M-2C	1780	11.8	0.0066
CYP2C8	1520	2.5	0.0017
CYP2C9	335	16.2	0.048
CYP2C19	649	32.4	0.050
Liver microsomal fraction[b]			
HLM	318 (K_{m1})	185 (V_{max1})	0.582 (V_{max1}/K_{m1})
	72.7 (K_{m2})	246 (V_{max2})	3.38 (V_{max2}/K_{m2})
MLM	1170	470	0.402

[a] V_{max}, pmol/min/pmol CYP; V_{max}/K_m, μl/min/pmol CYP.
[b] V_{max}, pmol/min/mg protein; V_{max}/K_m, μl/min/mg protein. HLM, human liver microsomes; MLM, marmoset liver microsomes. Each value represents the mean of two independent determinations.

3.3. *Drug oxidation activities*

The recombinant P450 M-2C did not show any detectable oxidation activities towards PT or S-MP under the conditions used (Fig. 4A and D). A slight activity was observed for DF 4′-hydroxylation by P450 M-2C at the highest substrate concentration used (100 μM) (Fig. 4B). P450 M-2C exerted considerable TB *p*-methylhydroxylase activities, which were 20–50% those of CYP2C9 at substrate concentrations from 0.25 to 2.5 mM (Fig. 4C). Based on these results, we performed kinetic studies for TB *p*-methylhydroxylation by P450 M-2C and compared the results with those of human CYP2C8, CYP2C9 and CYP2C19.

TB *p*-methylhydroxylation by four recombinant CYP enzymes showed monophasic kinetics in Michaelis–Menten plots (data not shown). The kinetic parameters obtained are summarized in Table 2. The recombinant CYP enzymes could be divided into two groups, i.e., high-K_m group (P450 M-2C and CYP2C8) and low-K_m group (CYP2C9 and CYP2C19).

TB *p*-methylhydroxylation by microsomal fractions from human and marmoset livers was analyzed by Eadie–Hofstee plots (data not shown). In human liver, microsomal TB

Table 3 – The IC_{50} values for inhibitors of TB *p*-methylhydroxylation by marmoset liver microsomes and recombinant P450 M-2C

Inhibitor	Marmoset liver microsomes		Recombinant P450 M-2C	
	0.1 mM[a]	1 mM[a]	1 mM[a]	2 mM[a]
Quercetin	51.4	16.1	105	61.2
Sulfaphenazole	>200	>200	>200	>200
Omeprazole	146	247	328	274

[a] Substrate concentration. IC_{50} values are expressed as μM.

p-methylhydroxylation showing biphasic kinetics, the higher K_m value (K_{m1}, 320 μM) was close to that of CYP2C9 (340 μM), while the lower K_m value (K_{m2}, 70 μM) was much smaller than any K_m values of the recombinant CYP enzymes examined (Table 2). This indicates that together with CYP2C9, another CYP enzyme having a lower K_m value is also involved in TB p-methylhydroxylation by the pooled human liver microsomal fractions employed. On the other hand, in marmoset liver microsomal TB oxidation showing monophasic kinetics (data not shown), the K_m value (1.2 mM) was close to that of P450 M-2C (1.8 mM) (Table 2), suggesting that P450 M-2C is the major TB p-methylhydroxylase in the mormoset liver.

3.4. *Inhibition studies*

The effects of three kinds of inhibitors, quercetin as a CYP2C8 inhibitor [25], sulfaphenazole as a CYP2C9 inhibitor [26] and omeprazole as a CYP2C19 inhibitor [27], on TB p-methylhydroxylation by microsomal fractions from mormoset liver (Fig. 5, upper panels) and yeast cells expressing P450 M-2C (Fig. 5, lower panels) were examined using two substrate concentrations of 0.1 and 1 mM. Quercetin (Fig. 5A and D) and omeprazole (Fig. 5C and F) similarly inhibited the TB oxidation activity of marmoset liver microsomes and recombinant P450 M-2C in a concentration-dependent manner. The potency of sulfaphenazole was lower than those of the other inhibitors (Fig. 5B and E). Table 3 lists the IC_{50} values for the inhibitors. The potencies of the inhibitors were ranked as quercetin > omeprazole > sulfaphenazole for both marmoset liver microsomes and recombinant P450 M-2C.

4. Discussion

In the present study, we have cloned a cDNA encoding a novel CYP enzyme from the fresh liver of an adult female marmoset. The deduced amino acid sequence exhibited high identities to human CYP2C8 (87%), crab-eating monkey CYP2C20 (89%) and rhesus monkey CYP2C74 (89%). The nucleotide and amino acid sequences were registered to GenBank (accession no. AB242600). Dr. David Nelson, University of Tennessee Menphis, recommended us to call this CYP "marmoset CYP2C8" (his personal communication). In this paper, however, we tentatively called the enzyme P450 M-2C, standing for the Marmoset CYP2C enzyme to avoid confusion with human CYP2C8.

According to the list of P450 families and subfamilies of Dr. Nelson's home page (http://drnelson.utmem.edu/P450.stats.all.2005.htm), four monkey cDNA sequences encoding CYP2C enzymes had been registered as of January 8, 2005: crab-eating monkey CYP2C20 (S53046), rhesus monkey CYP2C43 (AB212264), CYP2C74 (AY635462) and CYP2C75 (AY635463). The functions of these monkey CYP enzymes have not been studied in detail, except for CYP2C43.

Matsunaga et al. [28] cloned a cDNA encoding CYP2C43 and characterized the enzymatic properties of CYP2C43 protein expressed in yeast cells. They reported that the recombinant CYP2C43 catalyzed S-MP 4′-hydroxylation but not TB p-methylhydroxylation under the conditions they employed. Interestingly, marmoset P450 M-2C showed the reverse substrate specificity, i.e., it catalyzed TB p-methylhydroxylation but not S-MP 4′-hydroxylation.

P450 M-2C showed considerable oxidation activity only for TB among the four substrates of human CYP2C enzymes examined. Although all of the human CYP2C enzymes (CYP2C8, CYP2C9 and CYP2C19) examined exerted TB oxidation activities, the kinetic profile of CYP2C8 was most similar to that of marmoset P450 M-2C (Table 2). The results of the inhibition study demonstrated that quercetin, a CYP2C8 inhibitor, was the most effective inhibitor for TB oxidation by P450 M-2C as well as by marmoset liver microsomes, followed by omeprazole, a CYP2C19 inhibitor. TB p-methylhydroxylation was kinetically

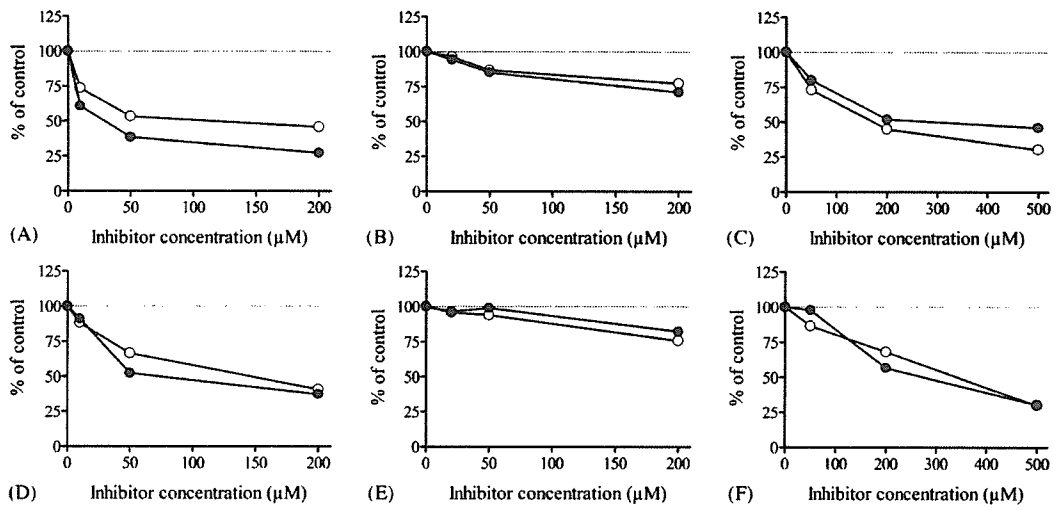

Fig. 5 – The effects of human CYP2C enzyme inhibitors on TB p-methylhydroxylation by marmoset liver microsomes (upper panels) and by P450 M-2C (lower panels). The final inhibitor concentrations used were 10, 50 and 200 μM for quercetin (A and D) 20, 50 and 200 μM for sulfaphenazole (B and E) and 50, 200 and 500 μM for omeprazole (C and F). The substrate concentrations used were 100 (open circles) and 1000 μM (closed circles). Each point represents the mean of two independent determinations.

analyzed to be monophasic, and the apparent K_m values were similar between the marmoset liver microsomes and the recombinant P450 M-2C, indicating that P450 M-2C is the major TB *p*-methylhydroxylase in the marmoset liver.

It is well known that CYP2C9 is the major TB *p*-methylhydroxylase in the human liver [26]. However, TB *p*-methylhydroxylation gave biphasic kinetics in the pooled human liver microsomes used in the present study. The apparent K_m value for TB *p*-methylhydroxylation by recombinant CYP2C9 was 340 μM in this study, which was close to the K_m values of purified CYP2C9 reported by Lasker et al. [29] (180–400 μM) and of recombinant CYP2C9 (410 μM) reported by Flanagan et al. [30] for TB *p*-methylhydroxylation. Therefore, it is reasonable to think that some CYP enzyme(s) with a lower K_m value of around 70 μM together with CYP2C9 with a higher K_m value of 340 μM are responsible for TB *p*-methylhydroxylation in the human liver microsomal fractions used.

As described above, for TB *p*-methylhydroxylation, P450 M-2C and CYP2C8 showed similar kinetic profiles in the present study. In contrast, P450 M-2C did not show any detectable activity for PT 6α-hydroxylation, which was catalyzed by CYP2C8. Fig. 6 shows the active sites of P450 M-2C and CYP2C8. In a modeling study on PT 6α-hydroxylation by CYP2C8, Tanaka et al. [31] proposed that there are two distal sites (1 and 2) in addition to the proximal site occupying the space just above the heme iron in the active site of CYP2C8. They thought

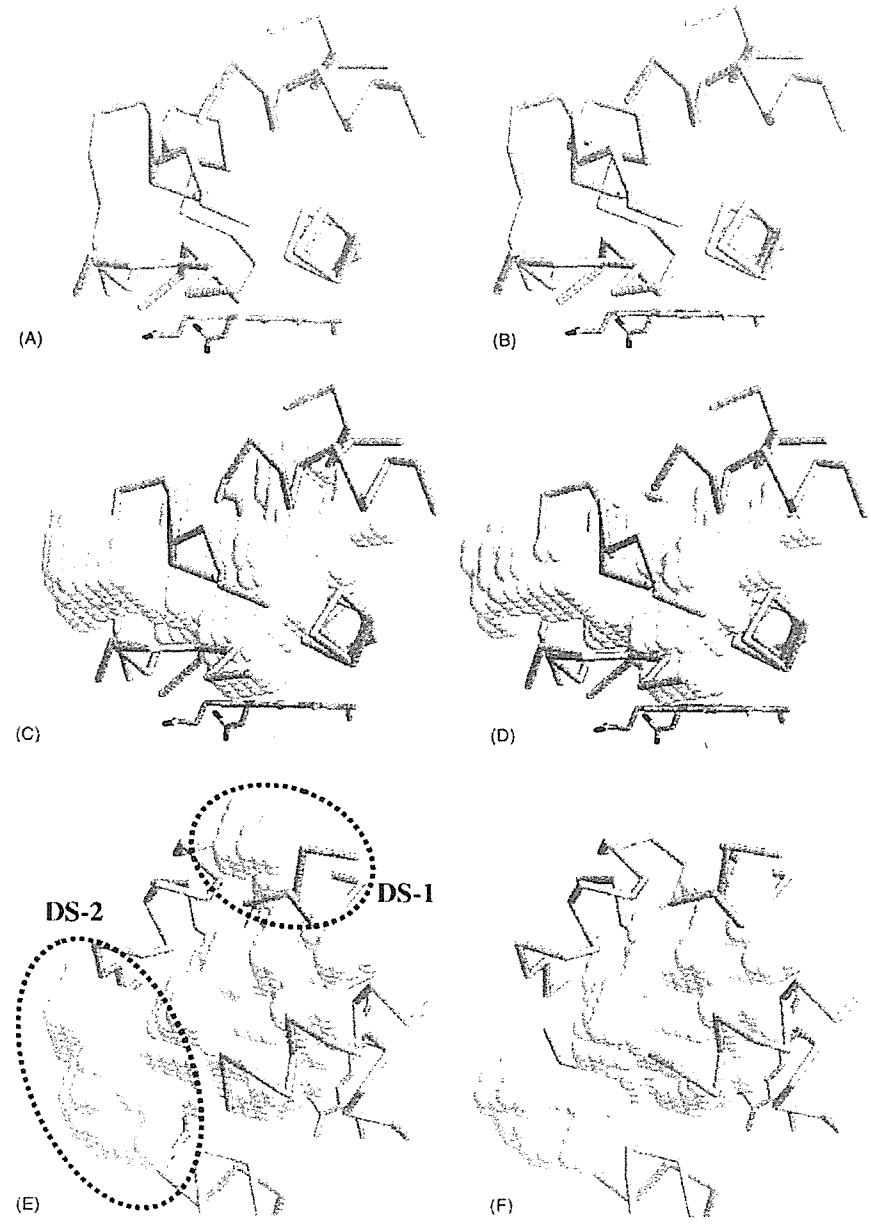

Fig. 6 – Comparison of the active site structures (A and B) and the active site cavities (C–F) between CYP2C8 (left panels) and P450 M-2C (right panels). The active site conformation was depicted using RasMol Version 2.6-ucb 1.0. DS, distal site.

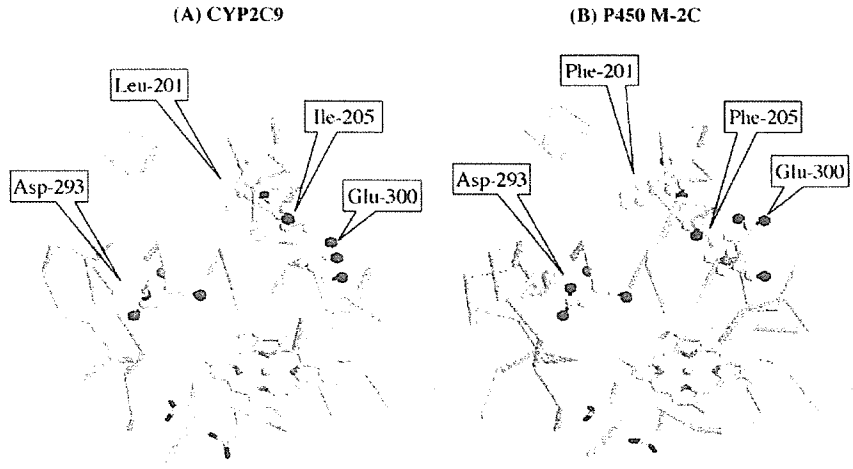

Fig. 7 – Comparison of the active site structures between CYP2C9 (A) and P450 M-2C (B). Proteins are depicted as backbone form, and Phe-201 and Asp-293 for CYP2C8 and Phe-205 and Glu-300 for P450 M-2C as ball and stick form using RasMol Version 2.6-ucb 1.0.

it possible that the N-benzoyl-3-phenylisoserine side-chain of PT binds to distal site 2, resulting in oxidation at the 6-position of the taxol ring.

The active site conformations of CYP2C8 and P450 M-2C which are shown with backbone depiction on RasMol are very similar (Fig. 6A and B). However, the shapes of the active site cavities of the two enzymes are considerably different from each other (Fig. 6C–F), especially, the shapes viewed from above (Fig. 6E and F) show clear differences in both distal sites 1 and 2. That is, the sizes of both distal sites 1 and 2 of CYP2C8 are larger than those of P450 M-2C. It is feasible that the smaller size of the active site cavity, particularly of the distal site 2 of P450 M-2C, makes it impossible for PT to appropriately dock in the active site, resulting in undetectable PT oxidation activity.

Fig. 7 shows the active sites of CYP2C9 (left panel) and P450 M-2C (right panel). The active site of CYP2C8 is almost the same as that of P450 M-2C. In the active site cavity of CYP2C9, there are two acidic amino acids, i.e., Asp-293 and Glu-300, whose carboxylate groups may interact ionically with basic nitrogen atoms of TB. As a result, the *p*-methyl group of TB to be oxidized comes close to the heme iron, yielding *p*-hydroxymethyl-TB efficiently. In the active site cavity of P450 M-2C as well as of CYP2C8 having Asp-293 and Glu-300, however, there are two aromatic amino acids, Phe-201 and Phe-205. The phenyl group of Phe-205, in particular, is located just in front of the carboxylate group of Glu-300, which seems to block the ionic interaction between Glu-300 and the basic nitrogen atom of TB. Furthermore, these phenylalanine residues may cause hydrophobic interaction with the aromatic ring of TB, making the tolyl group of TB far from the heme ion, which may result in low capacities of P450 M-2C and CYP2C8 for TB *p*-methylhydroxylation.

In summary, we cloned a cDNA encoding a novel CYP2C enzyme, called P450 M-2C, from the marmoset liver. The deduced amino acid sequence showed high identities to human CYP2C8 (87%), CYP2C9 (78%) and CYP2C19 (77%). Yeast cell microsomal P450 M-2C catalyzed *p*-methylhydroxylation of only TB among four substrates, PT, DF, TB and S-MP, for human CYP2C enzymes. Marmoset liver microsomes exerted monophasic kinetics for TB, and its apparent K_m value was similar to that of the recombinant P450 M-2C. Although three human recombinant CYP2C enzymes, CYP2C8, CYP2C9 and CYP2C19, also showed TB *p*-methylhydroxylation, the kinetic profile of CYP2C8 was most similar to that of P450 M-2C. TB oxidation by the marmoset microsomes and the recombinant P450 M-2C was similarly inhibited by quercetin, a CYP2C8 inhibitor. These results indicate that P450 M-2C (marmoset CYP2C8) is the major TB *p*-methylhydroxylase in the marmoset liver.

Acknowledgments

This study was financially supported by a grant-in-aid for scientific research from the Ministry of Education, Science, Culture and Sports of Japan. We thank Miss Tomoko Sumada for her technical assistance.

REFERENCES

[1] Rendic S, Di Carlo FJ. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers and inhibitors. Drug Metab Rev 1997;29:413–580.
[2] Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. Science 1999;286:487–91.
[3] Parkinson A, Mudra DR, Johnson C, Dwyer A, Carroll M. The effects of gender, age, ethnicity, and liver cirrhosis on cytochrome P450 enzyme activity in human liver microsomes and inducibility in cultured human hepatocytes. Toxicol Appl Pharmacol 2004;199:193–209.
[4] Kastner M, Schulz-Schalge T, Neubert D. Purification and properties of cytochrome P-450 from liver microsomes of phenobarbital-treated marmoset monkeys (*Callithrix jacchus*). Toxicol Lett 1989;45:261–70.

[5] Sakuma T, Igarashi T, Hieda M, Ohgiya S, Isogai M, Ninomiya S, et al. Marmoset CYP1A2: primary structure and constitutive expression in livers. Carcinogenesis 1997;18:1985–91.
[6] Edwards RJ, Murray BP, Murray S, Schulz T, Neubert D, Gant TW, et al. Contribution of CYP1A1 and CYP1A2 to the activation of heterocyclic amines in monkeys and human. Carcinogenesis 1994;15:829–36.
[7] Schulz TG, Neubert D, Davies DS, Edwards RJ. Inducibility of cytochromes P-450 by dioxin in liver and extrahepatic tissues of the marmoset monkey (*Callithrix jacchus*). Biochim Biophys Acta 1996;1298:131–40.
[8] Schulz TG, Thiel R, Davies DS, Edwards RJ. Identification of CYP2E1 in marmoset monkey. Biochim Biophys Acta 1998;1382:287–94.
[9] Schulz TG, Thiel R, Neubert D, Brassil PJ, Schulz-Utermoehl T, Boobis AR, et al. Assessment of P450 induction in the marmoset monkey using targeted anti-peptide antibodies. Biochim Biophys Acta 2001;1546:143–55.
[10] Igarashi T, Sakuma T, Isogai M, Nagata R, Kamataki T. Marmoset liver cytochrome P450s: study for expression and molecular cloning of their cDNAs. Arch Biochem Biophys 1997;339:85–91.
[11] Narimatsu S, Oda M, Hichiya H, Isobe T, Asaoka K, Hanioka N, et al. Molecular cloning and functional analysis of cytochrome P450 1A2 from Japanese monkey liver: comparison with marmoset cytochrome P450 1A2. Chemico-biol Interact 2005;152:1–12.
[12] Hichiya H, Kuramoto S, Yamamoto S, Shinoda S, Hanioka N, Narimatsu S et al. Cloning and functional expression of a novel marmoset cytochrome P450 2D enzyme, CYP2D30: comparison with the known marmoset CYP2D19. Biochem Pharmacol 2004; 68:165–75.
[13] Narimatsu S, Yonemoto R, Saito K, Takaya K, Kumamoto T, Ishikawa T, et al. Oxidative metabolism of 5-methoxy-*N*,*N*-diisopropyltryptamine (Foxy) by human liver microsomes and recombinant cytochrome P450 enzymes. Biochem Pharmacol 2006;71:1377–85.
[14] Wan J, Imaoka S, Chow T, Hiroi T, Yabusaki Y, Funae Y. Expression of four rat CYP2D isoforms in *Saccharomyces cerevisiae* and their catalytic specificity. Arch Biochem Biophys 1997;348:383–90.
[15] Tsuzuki D, Takemi C, Yamamoto S, Tamagake K, Imaoka S, Funae Y, et al. Functional evaluation of cytochrome P450 2D6 with Gly42Arg substitution expressed in *Saccharomyces cerevisiae*. Pharmacogenetics 2001;11:709–18.
[16] Omura T, Sato R. The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. J Biol Chem 1964;239:2370–8.
[17] Narimatsu S, Gotoh M, Masubuchi Y, Horie Y, Ohmori S, Kitada M, et al. Stereoselectivity in bunitrolol 4-hydroxylation in liver microsomes from marmosets and Japanese monkeys. Biol Pharm Bull 1996;19:1429–33.
[18] Guengerich FP, Wang P, Davidson NK. Estimation of isozymes of microsomal cytochrome P-450 in rats, rabbits, and humans using immunochemical staining coupled with sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Biochemistry 1982;21:1698–706.
[19] Soyama A, Saito Y, Hanioka N, Murayama N, Nakajima O, Katori N, et al. Non-synonumous single nucleotide alterations found in the CYP2C8 gene result in reduced in vitro paclitaxel metabolism. Biol Pharm Bull 2001;24: 1427–30.
[20] Schmitz G, Lepper H, Estler CJ. High-performance liquid chromatographic method for the routine determination of diclofenac and its hydroxy and methoxy metabolites from *in vitro* systems. J Chromatogr 1993;620:158–63.
[21] Komatsu K, Ito K, Nakajima Y, Kanamitsu S, Imaoka S, Funae Y, et al. Prediction of *in vivo* drug-drug interactions between tolbutamide and various sulfonamides in humans based on in vitro experiments. Drug Metab Dispos 2000;28:475–81.
[22] Nakajima M, Inoue T, Shimada N, Tokudome S, Yamamoto T, Kuroiwa Y. Cytochrome P450 2C9 catalyzes indomethacin O-demethylation in human liver microsomes. Drug Metab Dispos 1998;26:261–6.
[23] Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. Protein measurement with the Folin phenol reagent. J Biol Chem 1951;193:265–75.
[24] Masuda K, Tamagake K, Katsu T, Torigoe F, Saito K, Hanioka N, et al. The roles of phenylalanine at position 120 and glutamic acid at position 222 in the oxidation of chiral substrates by cytochrome P450 2D6. Chirality 2006;18:167–76.
[25] Rahman A, Korzekwa KR, Grogan J, Gonzalez FJ, Harris JW. Selective biotransformation of taxol to 6α-hydroxytaxol by human cytochrome P450 2C8. Cancer Res 1994;54:5543–6.
[26] Mancy A, Dijols S, Poli S, Guengerich P, Mansuy D. Interaction of sulfaphenazole derivatives with human liver cytochromes P450 2C: molecular origin of the specific inhibitory effects of sulfaphenazole on CYP 2C9 and consequences for the substrate binding site topology of CYP 2C9. Biochemistry 1996;35:16205–12.
[27] Ko JW, Sukhova N, Thacker D, Chen P, Flockhart DA. Evaluation of omeprazole and lansoprazole as inhibitors of cytochrome P450 isoforms. Drug Metab Dispos 1997;25:853–62.
[28] Matsunaga T, Ohmori S, Ishida M, Sakamoto Y, Nakasa H, Kitada M. Molecular cloning of monkey CYP2C43 cDNA and expression in yeast. Drug Metab Pharmacokinet 2002;17:117–24.
[29] Lasker JM, Wester MR, Aramsombatdee E, Raucy JL. Characterization of CYP2C19 and CYP2C9 from human liver: respective roles in microsomal tolbutamide, S-mephenytoin, and omeprazole hydroxylations. Arch Biochem Biophys 1998;353:16–28.
[30] Flanagan JU, McLaughlin LA, Paine MJ, Sutcliffe MJ, Roberts GC, Wolf CR. Role of conserved Asp293 of cytochrome P450 2C9 in substrate recognition and catalytic activity. Biochem J 2003;370:921–6.
[31] Tanaka T, Kamiguchi N, Okuda T, Yamamoto Y. Characterization of the CYP2C8 active site by homology modeling. Chem Pharm Bull 2004;52:836–41.

Drug Metab. Pharmacokinet. 21 (4): *SNP30* (337)–*SNP39* (346) (2006).

SNP Communication

Novel Genetic Variations and Haplotypes of Hepatocyte Nuclear Factor 4α (HNF4A) Found in Japanese Type II Diabetic Patients

Hiromi FUKUSHIMA-UESAKA[1,2], Yoshiro SAITO[1,2,*], Keiko MAEKAWA[1,2], Mayumi SAEKI[1], Naoyuki KAMATANI[3], Hiroshi KAJIO[4], Nobuaki KUZUYA[4], Kazuki YASUDA[5] and Jun-ichi SAWADA[1,2]

[1]*Project Team for Pharmacogenetics,* [2]*Division of Biochemistry and Immunochemistry, National Institute of Health Sciences, Tokyo, Japan;* [3]*Division of Genomic Medicine, Department of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo, Japan;* [4]*Division of Endocrine and Metabolic Diseases, the Hospital,* [5]*Department of Metabolic Disorder, Research Institute, International Medical Center of Japan, Tokyo, Japan*

Full text of this paper is available at http://www.jstage.jst.go.jp/browse/dmpk

Summary: Thirty-nine single nucleotide variations, including 16 novel ones, were found in the 5′ promoter region, all of the exons and their surrounding introns of *HNF4A* in 74 Japanese type II diabetic patients. The following novel variations were identified (based on the amino acid numbering of splicing variant 2): −208G>C in the 5′ promoter region; 1154C>T (A385V) and 1193T>C (M398T) in the coding exons; 1580G>A, 1852G>T, 2180C>T, 2190G>A, and 2362_2380delAAGAATGGTGTGGGAGAGG in the 3′-untranslated region, and IVS1+231G>A, IVS2−83C>T, IVS3+50C>T, IVS3−54delC, IVS5+173_176delTTAG, IVS5−181_−180delAT, IVS8−106A>G, and IVS9−151A>C in the introns. The allele frequencies were 0.311 for 2362_2380delAAGAATGGTGTGGGAGAGG, 0.054 for 1580G>A, 0.047 for 1852G>T, 0.020 for IVS1+231G>A, 0.014 for IVS9−151A>C, and 0.007 for the other 11 variations. In addition, one known nonsynonymous single nucleotide polymorphism, 416C>T (T139I), was detected at a 0.007 frequency. Based on the linkage disequilibrium profiles, the region analyzed was divided into three blocks. Haplotype analysis determined/inferred 10, 16, and 12 haplotypes for block 1, 2, and 3, respectively. Our results on *HNF4A* variations and haplotypes would be useful for pharmacogenetic studies in Japanese.

Key words: *HNF4A*; genetic variation; amino acid alteration; haplotype

Introduction

Hepatocyte nuclear factor 4α is an orphan nuclear receptor. This transcriptional factor is predominantly expressed in the liver, small intestine, colon, kidney, and pancreas and acts as a homodimer.[1] To date, nine transcript variants have been reported with alternative initiation and splicing of four exon 1's, and alternative splicing in exons 8–10.[2] As for the exon 1's, 6 isoforms (isoforms 1–6) are transcribed from exon 1A with the P1 promoter. Isoforms 4–6 use additional exon 1B, which confer an extra 30 amino acids as compared with the corresponding isoforms 1–3. The transcription of 3 isoforms (isoforms 7–9) starts from exon 1D with the P2 promoter approximately 45 kb upstream of exon 1A.[3] Isoforms 2, 5, and 8 use the alternative splice donor site in exon 9, 30 bases downstream of the site of isoforms 1, 4, and 7, and thus are alternatively spliced forms of isoforms 1, 4, and 7. On the other hand, the isoforms 3, 6, and 9 utilize exon 8 with a 125-base extension,

On Dec. 8, 2005, novel variations were not reported in the databases of the Japanese Single Nucleotide Polymorphisms (JSNP) (http://snp.ims.u-tokyo.ac.jp/), dbSNP in the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/SNP/), or PharmGKB (http://www.pharmgkb.org/do/).

Received; December 12, 2005, Accepted; March 9, 2006

**To whom correspondence should be addressed*: Yoshiro SAITO, Ph.D., *Division of Biochemistry and Immunochemistry, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. Tel. +81-3-5717-3831, Fax. +81-3-5717-3832, E-mail: yoshiro@nihs.go.jp*

resulting in an early stop codon in exon 8.[4] The mRNA level of isoform 1 or 2 was reported to be higher than that of isoform 3 in the liver.[4] Variant 4, and probably 5 and 6, has no detectable transactivation potential.[1] The transcription of HNF4α mainly initiates at the P1 promoter in adult liver and kidney.[2,3] In the pancreas, the transcription starts from the P2 promoter,[2,3] though one report showed mRNA expression of isoforms 1 and 2 in pancreatic islets and β cells.[5]

Human HNF4α protein can be divided into 6 domains named domains A (the N-terminal domain) to F (the C-terminal domain).[1,6] Domains A and B (amino acid residues 1–60 based on isoform 2) contain a ligand-independent transactivation function AF-1. Domains C (61–126) and D (127–143) are important for DNA-binding and full transcriptional activity, respectively. Domain E (144–378) contains a ligand-binding domain and a ligand-dependent transactivation function AF-2. The domains D and E are also important for functional interactions with co-activators such as PGC-1 and SRC-3.[7] Domain F (379–474) has been shown to inhibit the transactivation potential of AF-2.[6] In addition, it was reported that isoform 8 showed markedly reduced transcriptional activity compared to isoform 2, probably due to lack of AF-1 activity by usage of exon 1D, instead of exon 1A.[5,8]

Nuclear factor HNF4α has been reported to be involved in the induction of many drug metabolizing enzymes, such as *CYP2C9*, *CYP2C8*, *CYP2A6*, *CYP3A4*, and *UGT1A9* with known binding sites.[9–13] Interindividual differences in mRNA, protein, and activity levels have been shown in these drug metabolizing enzymes, and several genetic variations with functional significance have been found. Presently, however, the reported variations in these genes do not fully explain the interindividual differences. It may be possible that genetic variations of *HNF4A* may contribute to these differences. In the present study, the 5′-regulatory region, all the exons (except for non-functional exons 1B and 1C, and pancreatic exon 1D), and their surrounding introns of *HNF4A* were sequenced in 74 Japanese patients. Sixteen novel variations, including two nonsynonymous ones, were identified.

Materials and Methods

Human genomic DNA samples: DNA was extracted from the blood leukocytes of 74 Japanese type II diabetic patients who had received glimepiride. The ethical review boards of the International Medical Center of Japan and the National Institute of Health Sciences approved this study. Written informed consent was obtained from all participating patients.

Polymerase chain reaction (PCR) conditions for DNA sequencing: First, multiplex PCR was performed to amplify the entire *HNF4A* gene with two mixed primer sets (Mix 1 and Mix 2 in "1st PCR" in **Table 1**). Amplification was performed from 50 ng of genomic DNA using 1 unit of Ex-Taq (Takara Bio. Inc, Shiga, Japan) with 0.2 μM of the mixed primers sets. The first PCR conditions were 94°C for 5 min, followed by 30 cycles of 94°C for 30 sec, 60°C for 1 min, and 72°C for 2 min, and then a final extension at 72°C for 7 min. After the PCR products were treated with a PCR Product Pre-Sequencing Kit (USB Co., Cleveland, OH, USA), each exon was amplified separately using one-fifth of the volume of the 1st PCR product as a template by Ex-Taq (0.1 units) with a set of primers (0.2 μM) listed in "2nd PCR" of **Table 1** (designed in the intronic regions, except for exon 10, which is described below). The second-round PCR conditions were 94°C for 5 min, followed by 30 cycles of 94°C for 30 sec, 60°C for 1 min, and 72°C for 2 min, and then a final extension at 72°C for 7 min. Next, the PCR products were treated with a PCR Product Pre-Sequencing Kit and directly sequenced on both strands using an ABI BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) with the primers listed in "Sequencing" of **Table 1**. The excess dye was removed by a DyeEx96 kit (Qiagen, Hilden, Germany). The eluates were analyzed on an ABI Prism 3730 DNA Analyzer (Applied Biosystems). All detected rare variations were confirmed by repeating the PCR from the genomic DNA and sequencing the newly generated PCR products. Since exon 10 spans approximately 1.9 kb, the exon was amplified in two fragments in the 2nd PCR and sequenced with five forward and four reverse primers.

Linkage disequilibrium (LD) and haplotype analyses: Hardy-Weinberg equilibrium and LD analysis was performed by SNPAlyze software (Dynacom Co., Yokohama, Japan), and pairwise LD between variations was analyzed using r^2 and $|D'|$ values. For $|D'|$ values, the variations detected with a frequency greater than 0.05 were used. Based on the LD analysis, we divided the variations into 3 gene blocks and estimated haplotypes for each block. Some of the haplotypes were unambiguous from subjects with homozygous variations at all sites or a heterozygous variation at only one site. Separately, the diplotype configurations (a combination of haplotypes) were inferred by LDSUPPORT software, which determines the posterior probability distribution of the diplotype configuration for each subject based on the estimated haplotype frequencies.[14] The haplotypes inferred in single subjects are described with haplotype names and a question mark in **Tables 3 to 5**, since the predictability for these very rare haplotypes is known to be low in some cases. The haplotypes detected in this study were tentatively named as numbers plus small alphabetical letters. The block 2 **3a*

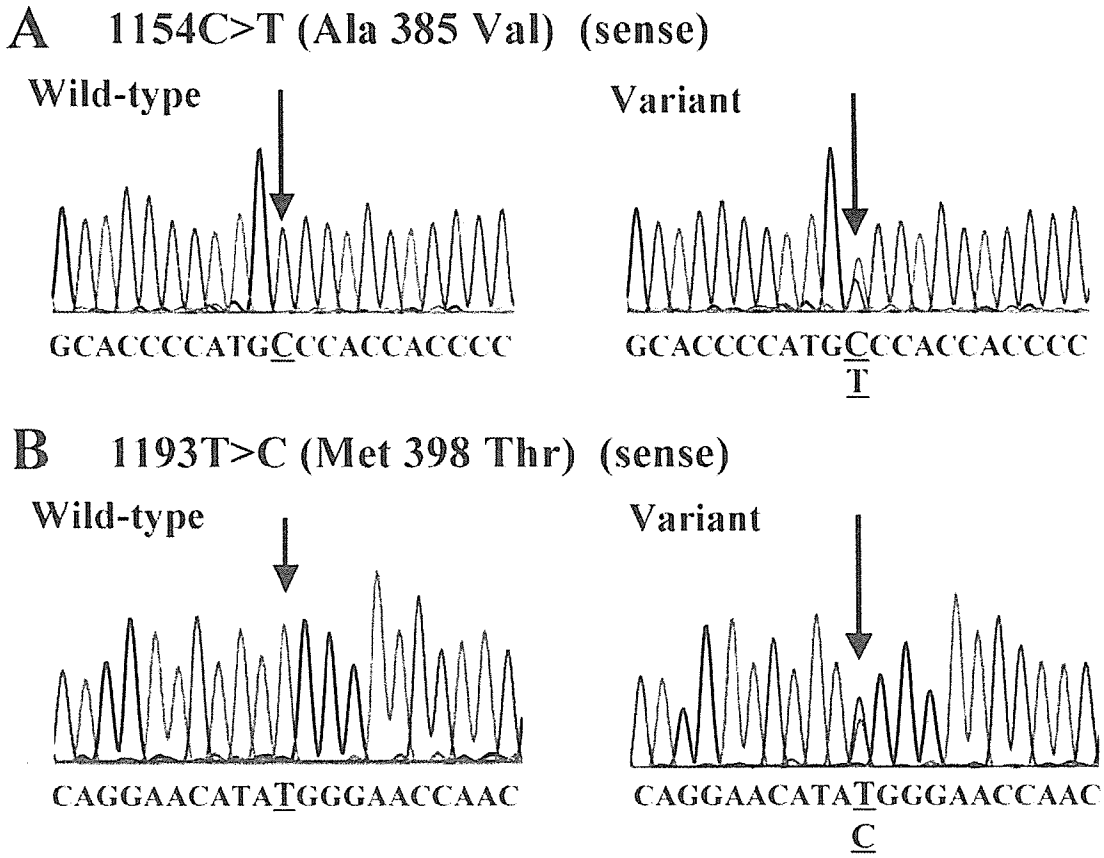

Fig. 1. Two novel nonsynonymous variations of human *HNF4A*.
(A) MPJ6_H4A_024 (wild-type, 1154C/C; variant, 1154C/T). (B) MPJ6_H4A_025 (wild-type, 1193T/T; variant, 1193T/C). Arrows indicate the positions of the nucleotide changes.

haplotype bearing 1154C>T (A385V) and 1193T>C (M398T) were confirmed as follows: the PCR fragment amplified by a high fidelity DNA polymerase KOD-Plus- (TOYOBO, Tokyo, Japan) with the primer pairs (5′-CACCGCACCTTGTTCCTTTCAACT-3′ as a forward primer and 5′-TGCCCTTTATTCCCTACCCT-3′ as a reverse primer) from the genomic DNA was cloned into pcDNA3.1-TOPO vector (Invitrogen, Carlsbad, CA, USA); and the cloned inserts (9 clones, *ca.*, 500 ng) were directly sequenced on both strands as described in the above section.

Results and Discussion

The P1 promoter regions (up to 300 bases upstream of the translational start site), all 10 exons (exon 1A and exons 2–10) and their flanking introns of *HNF4A* were sequenced in 74 Japanese type II diabetic patients. Genbank accession number NT_011362.9 was utilized for the reference sequence. The cDNA and amino acid numberings were based on isoform 2 of HNF4α (accession numbers NM_000457 and NP_000448, respectively, based on Drewes *et al.*[1]). Thirty-nine genetic variations, including 16 novel ones [1 in the promoter region, 2 in the coding exons, 5 in the 3′-untranslated region (3′-UTR), and 8 in the introns], were detected (see **Table 2**). All of the detected variations were found in Hardy-Weinberg equilibrium ($p>0.05$), except for IVS5−136T>C ($p=0.02$) because of a slightly low occurrence of homozygote compared with that of expectation.

Both of the two novel nonsynonymous variations, 1154C>T (A385V) and 1193T>C (M398T), were simultaneously found in a same patient as heterozygotes at 0.007 frequencies (**Fig. 1**), and their linkage was confirmed by cloning of the genomic DNA fragment and its sequencing analysis (data not shown). These variations are located in the F-domain of HNF4α, which has been shown to modulate the transactivation potential of AF-2.[9] According to PolyPhen, a prediction tool for the possible impact of amino acid substitutions (http://genetics.bwh.harvard.edu/pph/), it is possible that M398T (but not A385V) damages the function or structure of this protein. Moreover, a mutation located four amino acids downstream of

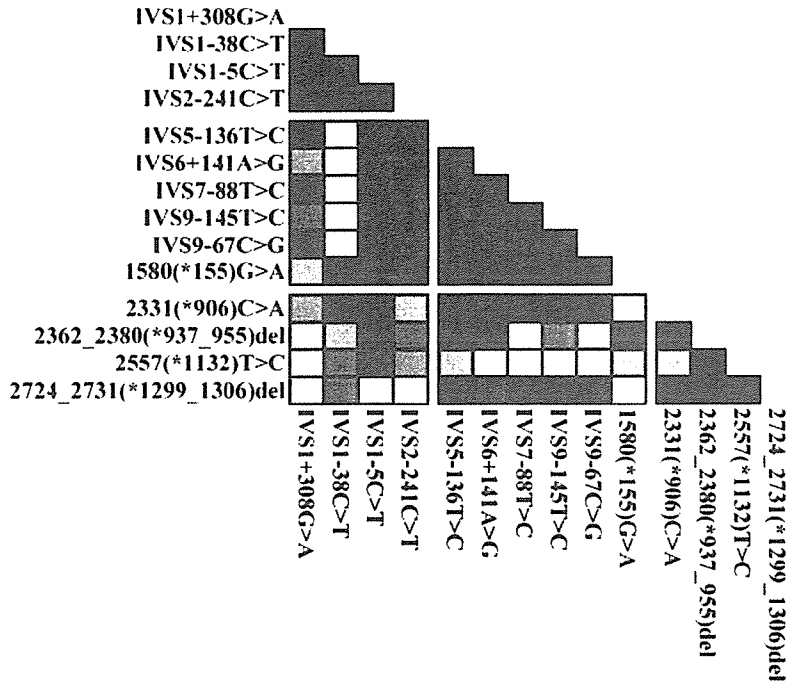

Fig. 2. Linkage disequilibrium (LD) analysis of *HNF4A* by |D'| statistics.
Pairwise LD (|D'|) between 14 common SNPs (> 0.05 in their allele frequencies) are shown by a 10-graded gray color. The denser color represents the higher linkage.

M398 (V402I mutation, V393I in the original paper) results in reduced transactivation activity.[15] Further functional analysis should be pursued for these variations.

HNF4α is also known as maturity-onset diabetes of young (MODY) gene, MODY-1. Many genetic variations have been found in MODY patients, and several in the Japanese. For example, R136W (R127W in the original paper), one of the mutations found in early (<25-years-old)-onset type II diabetic patients, was shown to decrease the DNA binding ability and transcriptional activity.[16,17] As for the newly identified nonsynonymous variations 1154C>T (A385V) and 1193T>C (M398T), the association with MODY is unlikely since the variations were found in a patient who was diagnosed with type II diabetes around 60 years of age.

The known nonsynonymous variation, 416C>T (T139I, T130I in the original paper), was shown to be associated with reduced transcriptional activity and type II diabetes in Japanese and Danish white subjects.[18,19] This variation was detected at a 0.007 frequency in our diabetic patients, which is lower than the reported frequency in Japanese diabetic patients (0.035 in 423 patients) but similar to that in the nondiabetic subjects (0.008 in 354 subjects).[18] The reason for this discrepancy is currently unknown, but this might be due to the small number of patients analyzed in the present study and relatively low frequency of this variation. The patient with this T139I heterozygous variation showed a good response to glimepiride. Her HbA1c value decreased from 8.3 to 6.9 by treatment with glimepiride (1 mg/day) for four months. It is possible that the reduced transcriptional activity of HNF4α from this variation might lead to reduced CYP2C9 levels, resulting in increased glimepiride bioavailability. The frequency of intronic variation IVS1−38C>T (0.196) was slightly lower than those in the previous report on Japanese early-onset type II diabetic patients (0.28) and nondiabetic subjects (0.24).[16] The frequency of IVS1−5C>T (0.223) was slightly higher than those (0.14 and 0.15) described in the previous reports for Japanese and Chinese nondiabetic subjects, respectively.[16,20,21] Instead, our frequency for IVS1−5C>T was similar to those of Japanese and Chinese type-II diabetic patients (0.216 and 0.24, respectively).[20,21]

Using the detected variations, linkage disequilibrium (LD) was first analyzed using r^2 values (data not shown). Although found only in two subjects, perfect linkages were observed among IVS3−204C>G, IVS4+140C>G, IVS4−197A>C, IVS4−96C>G, IVS4−52G>A, IVS6+196G>A, IVS9−151A>C, and 2331C>T ($r^2=1$). Because the novel nonsynonymous variations 1154C>T (A385V) and 1193T>C (M398T)

Table 1. Primer sequences used for the analysis of *HNF4A*

		Amplified or sequenced region	Forward primer (5′ to 3′)	Reverse primer (5′ to 3′)	Amplified length (bp)
1st PCR	Mix 1	Exon 1	GTCAAGGGTCAAATGAGTGC	CCTTGCCGTCTCTCTGAACC	902
		Exon 4	AGCCGAGTCTTCACTGTCTT	GAAGGTGAAGACTCTGCTTG	626
		Exon 6	CACACAGAATGTTGCTTACA	TCGTGCTCTGACTTCAATGC	633
		Exon 8	TCTTTTTCTGCCTGTGTCTA	ACTGAGGCACAGACAGGTTA	571
	Mix 2	Exon 2 to 3	TACAGATGTGAAACTGAAGC	CTCTTCTCAGCCATTAGCCA	1,733
		Exon 5	CAGACTCCTTGGGGCTCTAA	CACCTATGAGATGGCAGTAA	654
		Exon 7	GATGCTTTGGTGCCTATCAG	CTTGACTTGCCTCATCTGTT	617
		Exon 9	GCACCTTGTTCCTTTCAACT	TGCCCTTTATTCCCTACCCT	529
		Exon 10	GACCTTCCAGACCTCATAAA	GGGTCTAATGCTTCCAGAAT	2,307
2nd PCR		Exon 1	GTCAAGGGTCAAATGAGTGC	CCTTGCCGTCTCTCTGAACC	902
		Exon 2	TACAGATGTGAAACTGAAGC	AAGACTTAGTATTGTGCCTG	606
		Exon 3	TACTCCAGAGGTCAAGGTTC	CTCTTCTCAGCCATTAGCCA	522
		Exon 4	AGCCGAGTCTTCACTGTCTT	GAAGGTGAAGACTCTGCTTG	626
		Exon 5	CAGACTCCTTGGGGCTCTAA	CACCTATGAGATGGCAGTAA	654
		Exon 6	CACACAGAATGTTGCTTACA	TCGTGCTCTGACTTCAATGC	633
		Exon 7	GATGCTTTGGTGCCTATCAG	CTTGACTTGCCTCATCTGTT	617
		Exon 8	TCTTTTTCTGCCTGTGTCTA	ACTGAGGCACAGACAGGTTA	571
		Exon 9	GCACCTTGTTCCTTTCAACT	TGCCCTTTATTCCCTACCCT	529
		Exon 10	GACCTTCCAGACCTCATAAA	TGGAGGAGAAAGGCGTCTTC	1,043
			GAAACGATTCCCCCAGTCAT	GGGTCTAATGCTTCCAGAAT	1,372
Sequencing		Exon 1	CTGAACATCGGTGAGTTAGG	CCTTGCCGTCTCTCTGAACC	
		Exon 2	TACAGATGTGAAACTGAAGC	AAGACTTAGTATTGTGCCTG	
		Exon 3	TACTCCAGAGGTCAAGGTTC	CTCTTCTCAGCCATTAGCCA	
		Exon 4	AGCCGAGTCTTCACTGTCTT	GAAGGTGAAGACTCTGCTTG	
		Exon 5	CAGACTCCTTGGGGCTCTAA	CACCTATGAGATGGCAGTAA	
		Exon 6	CACACAGAATGTTGCTTACA	TCGTGCTCTGACTTCAATGC	
		Exon 7	GATGCTTTGGTGCCTATCAG	CTTGACTTGCCTCATCTGTT	
		Exon 8	TCTTTTTCTGCCTGTGTCTA	ACTGAGGCACAGACAGGTTA	
		Exon 9	GCACCTTGTTCCTTTCAACT	TGCCCTTTATTCCCTACCCT	
		Exon 10	GACCTTCCAGACCTCATAAA	TATCCAGAGCAGGGCGTCAA	
			AGAGCAGGAATGGGAAGGAT	TGGAGGAGAAAGGCGTCTTC	
			GAAACGATTCCCCCAGTCAT	AAGACAGTGCCTGGGAGTAA	
			CACATCAGAGTGACATCCAG	GGGTCTAATGCTTCCAGAAT	
			CCTAGAGATTGTTTTTGTTT		

were found in the same patient, they were statistically estimated to be linked with each other, and this was confirmed by cloning and sequencing analysis as described above. By the same reason, the pairs, −208G>C and IVS5+173_176delTTAG, and IVS3−54delC and IVS8−106A>G, might be linked closely. Relatively strong LD ($r^2 \geqq 0.5$) was observed between IVS1−5C>T and IVS2−241C>T, and among IVS5−136T>C, IVS6+141A>G, IVS7−88T>C, IVS9−145T>C, and IVS9−67C>G. The r^2 values were below 0.5 for the other pairs of variations.

We found that haplotypes without block partitioning were too diverse. Thus, block partitioning for haplotype analysis was performed based on |D′| values using the 14 common variations detected with a frequency greater than 0.05 (**Fig. 2**). All the |D′| values were 1.0 between six pairs of the first four variations from IVS1+308G>A to IVS2−241C>T. Among the next six variations from IVS5−136T>C to 1580(*155)G>A, 14 out of 15 pairs (93%) had |D′| values over 0.9. If the 2331(*906)C>A was included, the percentage of the pairs over 0.9 was reduced to 76% (16/21). Among the last 4 variations [from 2331(*906)C>A to 2724_2731(*1299_1306)delTCCTCCCT], 4 out of 6 pairs (67%) had |D′| values over 0.9. If the 1580(*155)G>A was included, the percentage was reduced to 50% (5/10). Thus, the *HNF4A* haplotypes were analyzed as three blocks divided between IVS2−241C>T and IVS5−136T>C and between 1580(*155)G>A and 2331(*906)C>A. The boundary of block 1 and block 2 for the minor variations was tentatively assigned between IVS3+50C>T and IVS3−204C>G because of the perfect linkage ($r^2=1$) of IVS3−204C>G with IVS6+196G>A and a long distance between exons 3 and 4 (6.2 kb). The boundary of block 2 and block 3 for the minor variations was also tentatively assigned between 1817(*392)T>G and 1852(*427)G>T because of a moderate linkage ($r^2=0.36$) between 1852(*427)G>T and 2724_2731 (*1299_1306)delTCCTCCCT. The block partitions were

Table 2. Summary of *HNF4A* variations detected in this study

SNP ID			Reference	Location	Position		Nucleotide change and flanking sequences (5′ to 3′)	Amino acid change	Number of subjects			Frequency
This Study	JSNP	dbSNP (NCBI)			NT_011362.9	From the translational initiation site or from the end of the nearest exon[b]			Wild-type	Hetero-zygote	Homo-zygote	
MPJ6_H4A_001[a]				Promoter	8082720	−208	CGGCGGGGGACC G/C ATTAACCATTAA		73	1	0	0.007
MPJ6_H4A_002[a]				Intron 1	8083273	IVS1+231	CCAGCAAAAGTC G/A ATCCCGGCTATT		71	3	0	0.020
MPJ6_H4A_003	IMS-JST006533	rs2071197		Intron 1	8083350	IVS1+308	TGGCACAGTGAC G/A TGATGGTGAGCT		18	45	11	0.453
MPJ6_H4A_004	IMS-JST006534	rs2071198		Intron 1	8083399	IVS1+357	ATTCAGCCCAGC A/T CGGCCCCTTCGT		70	4	0	0.027
MPJ6_H4A_005		rs736824	17	Intron 1	8087575	IVS1−38	CCTGACATTCTG C/T TCTTCCTGAAGC		48	23	3	0.196
MPJ6_H4A_006		rs745975	17, 21, 22	Intron 1	8087608	IVS1−5	CTTCTCTCCTGG C/T GCAGACACGTCC		42	31	1	0.223
MPJ6_H4A_007		rs6093976		Intron 2	8088695	IVS2−241	CAACCAGCAGAG C/T GACCCAGGACCA		48	25	1	0.182
MPJ6_H4A_008[a]				Intron 2	8088853	IVS2−83	AGGTTGGGGGGT C/T AACTGGATAGCC		73	1	0	0.007
MPJ6_H4A_009[a]				Intron 3	8089080	IVS3+50	CACCTGCACCCA C/T AGCTCCCCGACA		73	1	0	0.007
MPJ6_H4A_010		rs13041396		Intron 3	8095045	IVS3−204	CTTATAGCCTCT C/G CATTGTGTTGGG		72	2	0	0.014
MPJ6_H4A_011[a]				Intron 3	8095195	IVS3−54	CACAGACACCCC C/- ACCCCCTACTCC		73	1	0	0.007
MPJ6_H4A_012		rs1800961	21, 22	Exon 4	8095279	416	ACCGGATCAGCA C/T TCGAAGGTCAAG	T139I	73	1	0	0.007
MPJ6_H4A_013		rs11574738		Intron 4	8095495	IVS4+140	GGGACACTGAGT C/G CGGTTTCACATG		72	2	0	0.014
MPJ6_H4A_014		rs11574739		Intron 4	8095865	IVS4−197	ACAGGTGAAGGC A/C CAGAGGGAGCCC		72	2	0	0.014
MPJ6_H4A_015		rs3212194		Intron 4	8095966	IVS4−96	CCAGCCCCCTCC C/G CACATCTGATTC		72	2	0	0.014
MPJ6_H4A_016		rs3212195		Intron 4	8096010	IVS4−52	AGGGGACAGAGA G/A TGCGGGAGGGCC		72	2	0	0.014
MPJ6_H4A_017[a]				Intron 5	8096390_3	IVS5+173_176	ATATTAACTCAG TTAG/- CTCCTCCAACAA		73	1	0	0.007
MPJ6_H4A_018[a]				Intron 5	8099799_800	IVS5−181_−180	AGCTCTGAGCAC AT/- GTTCTTTCCCCT		73	1	0	0.007
MPJ6_H4A_019	IMS-JST162873	rs3212200		Intron 5	8099844	IVS5−136	GGGTAGTAGTTT T/C ATGATGCCCATT		37	36	1	0.257
MPJ6_H4A_020		rs6103731		Intron 6	8100208	IVS6+141	CAGGCTTGCATT A/G GAGGGCTCCAGG		33	34	7	0.324
MPJ6_H4A_021		rs11086925		Intron 6	8100263	IVS6+196	ATGCAAGGAAAT G/A TGGATGCAAGTC		72	2	0	0.014
MPJ6_H4A_022	IMS-JST069233	rs2273618		Intron 7	8105485	IVS7−88	CTCCTTGTGTGA T/C ACAAGTCAGGGG		23	42	9	0.405
MPJ6_H4A_023[a]				Intron 8	8109784	IVS8−106	CAGGCACTGCCA A/G TATTGGATGGGC		73	1	0	0.007
MPJ6_H4A_024[a]				Exon 9	8109914	1154	ATGCACCCCATG C/T CCACCACCCCCT	A385V	73	1	0	0.007
MPJ6_H4A_025[a]				Exon 9	8109953	1193	TGCAGGAACATA T/C GGGAACCAACGT	M398T	73	1	0	0.007
MPJ6_H4A_026[a]				Intron 9	8110927	IVS9−151	ACCTTAGGGATT A/C TCTGGTTTAATT		72	2	0	0.014
MPJ6_H4A_027	IMS-JST098122	rs3746574		Intron 9	8110933	IVS9−145	GGGATTATCTGG T/C TTAATTAATTCT		33	37	4	0.304
MPJ6_H4A_028	IMS-JST098123	rs3746575		Intron 9	8111011	IVS9−67	AACTTTCCCGGG C/G CTCTTCATTTAC		24	42	8	0.392
MPJ6_H4A_029	IMS-JST098124	rs3746576		Exon 10	8111106	1311	ACAGCCCTCACC G/T CCAGGTGGCTCA	P437P	73	1	0	0.007
MPJ6_H4A_030[a]				3′-UTR	8111375	1580 (*155)[c]	CATGGCCTAAGG G/A CCACATCCCACT		66	8	0	0.054
MPJ6_H4A_031		rs11086926		3′-UTR	8111612	1817 (*392)[c]	TCTCCTAGCCCC T/G GTCATGGTGTCC		69	5	0	0.034
MPJ6_H4A_032[a]				3′-UTR	8111647	1852 (*427)[c]	CTGTGAGGCTGG G/T TCCAATTGTGGC		67	7	0	0.047
MPJ6_H4A_033[a]				3′-UTR	8111975	2180 (*755)[c]	GAGAAAACAAAC C/T CAGGTTGGCGAC		73	1	0	0.007
MPJ6_H4A_034[a]				3′-UTR	8111985	2190 (*765)[c]	ACCCAGGTTGGC G/A ACTGCAACAGGA		73	1	0	0.007
MPJ6_H4A_035		rs3212210		3′-UTR	8112126	2331 (*906)[c]	GAGCCAAGGCCT C/A GTGGTAGTAAGA		25	40	9	0.392
MPJ6_H4A_036		rs3212210		3′-UTR	8112126	2331 (*906)[c]	GAGCCAAGGCCT C/T GTGGTAGTAAGA		72	2	0	0.014
MPJ6_H4A_037[a]				3′-UTR	8112157_75	2362_2380 (*937_955)[c]	CAAGAATTGAGG AAGAATGGTGTGGGAGAGG/- GATGATGAAGAG		35	32	7	0.311
MPJ6_H4A_038		rs6130615		3′-UTR	8112352	2557 (*1132)[c]	GATGATATAATG T/C GGGTGAGAGTAA		28	31	15	0.412
MPJ6_H4A_039	IMS-JST114749	rs3834658		3′-UTR	8112519_26	2724_2731 (*1299_1306)[c]	TTAATCCTCCCT TCCTCCCT/- ATTAACCTAGAG		57	16	1	0.122

[a] Novel variations detected in this study.

[b] Exon-intron boundary and amino acid numbering were based on the isoform 2.

[c] Numbered from the termination codon TGA.

Table 3. Haplotypes of *HNF4A* (Block 1)

Nucleotide change[a]			-208G>C	IVS1+ 231G>A	IVS1+ 308G>A	IVS1+ 357A>T	IVS1- 38C>T	IVS1- 5C>T	IVS2- 241C>T	IVS2- 83C>T	IVS3+50 C>T	Number	Frequency
Amino acid change													
Haplotypes[b, c, d]	*1	*1a										63	0.426
		*1b										27	0.182
		*1c										24	0.162
		*1d										18	0.122
		*1e										6	0.041
		*1f										4	0.027
		*1g										3	0.020
		*1h										1	0.007
		*1i ?										1	0.007
		*1j ?										1	0.007

[a] A of the translational start codon of *HNF4A* is numbered 1. NT_011362.9 was used as the reference sequence.
[b] Major allele, white; minor allele, gray.
[c] The haplotypes are described as numbers plus small alphabetical letters.
[d] The haplotypes inferred in only one subject are described with haplotype names and a question mark.

Table 4. Haplotypes of *HNF4A* (Block 2)

Nucleotide change[a]			IVS3 -204 C>G	IVS3 -54 del C	416 C>T	IVS4 +140 C>G	IVS4 -197 A>C	IVS4 -96 C>G	IVS4 -52 G>A	IVS5 +173 _176 del TTAG	IVS5 -181 _-180 del AT	IVS5 -136 T>C	IVS6 +141 A>G	IVS6 +196 G>A	IVS7 -88 T>C	IVS8 -106 A>G	1154 C>T	1193 T>C	IVS9 -151 A>C	IVS9 -145 T>C	IVS9 -67 C>G	1311 G>T	1580 (*155)[b] G>A	1817 (*392)[b] T>G	Number	Frequency
Amino acid change					T139I												A385V	M398T				P437P				
Haplotypes[c, d, e]	*1	*1a																							85	0.574
		*1b																							35	0.236
		*1c																							8	0.054
		*1d																							4	0.027
		*1e																							2	0.014
		*1f																							2	0.014
		*1g																							2	0.014
		*1h																							2	0.014
		*1i																							1	0.007
		*1j																							1	0.007
		*1k ?																							1	0.007
		*1l ?																							1	0.007
		*1m ?																							1	0.007
		*1n ?																							1	0.007
	*2	*2a ?																							1	0.007
	*3	*3a																							1	0.007

[a] A of the translational start codon of *HNF4A* is numbered 1. NT_011362.9 was used as the reference sequence.
[b] Numbered from the termination codon TAG.
[c] Major allele, white; minor allele, gray.
[d] The haplotypes are described as numbers plus small alphabetical letters.
[e] The haplotypes inferred in only one subject are described with haplotype names and a question mark.

Table 5. Haplotypes of *HNF4A* (Block 3)

Nucleotide change[a]			1852 (*427)[b] G>T	2180 (*755)[b] C>T	2190 (*765)[b] G>A	2331 (*906)[b] C>A	2331 (*906)[b] C>T	2362_2380 (*937_955)[b] del AAGAATGGTG TGGGAGAGG	2557 (*1132)[b] T>C	2724_2731 (*1299_1306)[b] del TCCTCCCT	Number	Frequency
Amino acid change												
Haplotypes[c,d,e]	*1	*1a									39	0.264
		*1b									36	0.243
		*1c									26	0.176
		*1d									12	0.081
		*1e									10	0.068
		*1f									7	0.047
		*1g									7	0.047
		*1h									6	0.041
		*1i									2	0.014
		*1j									1	0.007
		*1k ?									1	0.007
		*1l ?									1	0.007

[a] A of the translational start codon of *HNF4A* is numbered 1. NT_011362.9 was used as the reference sequence.
[b] Numbered from the termination codon TAG.
[c] Major allele, white; minor allele, gray.
[d] The haplotypes are described as numbers plus small alphabetical letters.
[e] The haplotypes inferred in only one subject are described with haplotype names and a question mark.

similar to those inferred in Chinese reported previously, except for the regions that we did not analyze.[22]

Haplotype analysis was then performed (**Tables 3 to 5**). The haplotypes assigned in this study are shown as numbers plus small alphabetical letters. As for block 1 spanning 6.4 kb from the 5′ promoter region to intron 3, seven haplotypes were first unambiguously assigned by homozygous variations at all sites (**1a* to **1d*) or a heterozygous variation at only one site (**1e*, **1f*, and **1h*). Separately, the diplotype configurations (a combination of haplotypes) for all 74 patients were estimated with over 0.99 certainty by LDSUPPORT software. The additionally inferred haplotypes were three **1* subtypes (**1g*, **1i*, and **1j*). In our separate experiment, the **1g* haplotype was unambiguously identified by the presence of a **1g* homozygote in a cell line MEG-01 (data not shown). The determined/inferred haplotypes were summarized in **Table 3**. The most frequent haplotype was **1a* (frequency: 0.426), followed by **1b* (0.182), **1c* (0.162), and **1d* (0.122). The frequencies of the other haplotypes were less than 0.1.

Block 2 spans 16.6 kb from intron 3 to the 3′-UTR. Five haplotypes were unambiguously assigned by homozygous variations at all sites (**1a*) or a heterozygous variation at only one site (**1b*, **1d*, **1i*, and **1j*). From analysis with the software, the diplotype configurations for all 74 patients were estimated with over 0.99 certainty, except for 1 patient inferred to be **1d/*1h*. The additionally inferred haplotypes were nine **1* subtypes (**1c*, **1e* to **1h*, and **1k* to **1n*), **2a* (with 416C>T, T139I), and **3a* (with 1154C>T, A385V; 1193T>C, M398T; confirmed by cloning and sequencing analysis). The summary of the determined/inferred haplotypes was shown in **Table 4**. The most frequent haplotype was **1a* (frequency: 0.574), followed by **1b* (0.236). The frequencies of the other haplotypes including **2a* and **3a* were less than 0.1.

Regarding block 3 including eight 3′-UTR variations, nine haplotypes were unambiguously assigned by homozygous variations at all sites (**1a* to **1c* and **1e*) or a heterozygous variation at only one site (**1d*, **1f*, and **1h* to **1j*). From analysis with the software, the diplotype configurations for all 74 patients were estimated with over 0.99 certainty, except for 8 patients (inferred as **1a/*1c* for 6 patients and **1b/*1d* for 2 patients). The additionally inferred haplotypes by the software were three **1* subtypes (**1g*, **1k*, and **1l*). The determined/inferred haplotypes were summarized in **Table 5**. The haplotypes with more than 0.1 frequency were **1a* (0.264), **1b* (0.243), and **1c* (0.176).

In conclusion, 39 genetic variations, including 16 novel ones, were detected in *HNF4A* in the Japanese patients. Using the detected variations, 10, 16, and 12 haplotypes were determined and/or inferred for block 1, 2, and 3, respectively. Our results on *HNF4A* variations and haplotypes would be useful for pharmacogenetic studies in Japanese.

Acknowledgements: We appreciate great help from Drs. Kei Fujimoto and Kanae Yasuda at Hiranuma Clinic. We also thank Ms. Chie Sudo for her secretarial assistance.

Hiromi Fukushima-Uesaka and Yoshiro Saito contributed equally to this work.

This study was supported in part by the Program for the Promotion of Fundamental Studies in Health Sciences and by the Health and Labor Sciences Research Grants (Research on Advanced Medical Technology) from the Ministry of Health, Labor and Welfare of Japan.

References

1) Drewes, T., Senkel, S., Holewa, B. and Ryffel, G. U.: Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. *Mol. Cell. Biol.* **16**: 925-931 (1996).
2) Hansen, S. K., Parrizas, M., Jensen, M. L., Pruhova, S., Ek, J., Boj, S. F., Johansen, A., Maestro, M. A., Rivera, F., Eiberg, H., Andel, M., Lebl, J., Pedersen, O., Ferrer, J. and Hansen T.: Genetic evidence that HNF-1alpha-dependent transcriptional control of HNF-4alpha is essential for human pancreatic beta cell function. *J. Clin. Invest.* **110**: 827-833 (2002).
3) Thomas, H., Jaschkowitz, K., Bulman, M., Frayling, T. M., Mitchell, S. M., Roosen, S., Lingott-Frieg, A., Tack, C. J., Ellard, S., Ryffel, G. U. and Hattersley, A. T.: A distant upstream promoter of the HNF-4alpha gene connects the transcription factors involved in maturity-onset diabetes of the young. *Hum. Mol. Genet.* **10**: 2089-2097 (2001).
4) Kritis, A. A., Argyrokastritis, A., Moschonas, N. K., Power, S., Katrakili, N., Zannis, V. I., Cereghini, S. and Talianidis, I.: Isolation and characterization of a third isoform of human hepatocyte nuclear factor 4. *Gene* **173**: 275-280 (1996).
5) Eeckhoute, J., Moerman, E., Bouckenooghe, T., Lukoviak, B., Pattou, F., Formstecher, P., Kerr-Conte, J., Vandewalle, B. and Laine, B.: Hepatocyte nuclear factor 4 alpha isoforms originated from the P1 promoter are expressed in human pancreatic beta-cells and exhibit stronger transcriptional potentials than P2 promoter-driven isoforms. *Endocrinology* **144**: 1686-1694 (2003).
6) Hadzopoulou-Cladaras, M., Kistanova, E., Evagelopoulou, C., Zeng, S., Cladaras, C. and Ladias, J. A.: Functional domains of the nuclear receptor hepatocyte nuclear factor 4. *J. Biol. Chem.* **272**: 539-550 (1997).
7) Iordanidou, P., Aggelidou, E., Demetriades, C. and Hadzopoulou-Cladaras, M.: Distinct amino acid residues may be involved in coactivator and ligand interactions in hepatocyte nuclear factor-4alpha. *J. Biol.*

Chem. **280**: 21810-21819 (2005).
8) Ihara, A., Yamagata, K., Nammo, T., Miura, A., Yuan, M., Tanaka, T., Sladek, F. M., Matsuzawa, Y., Miyagawa, J. and Shimomura, I.: Functional characterization of the HNF4alpha isoform (HNF4alpha8) expressed in pancreatic beta-cells. *Biochem. Biophys. Res. Commun.* **329**: 984-990 (2005).
9) Chen, Y., Kissling, G., Negishi, M. and Goldstein, J. A.: The nuclear receptors constitutive androstane receptor and pregnane X receptor cross-talk with hepatic nuclear factor 4alpha to synergistically activate the human CYP2C9 promoter. *J. Pharmacol. Exp. Ther.* **314**: 1125-1133 (2005).
10) Ferguson, S. S., Chen, Y., LeCluyse, E. L., Negishi, M. and Goldstein, J. A.: Human CYP2C8 is transcriptionally regulated by the nuclear receptors constitutive androstane receptor, pregnane X receptor, glucocorticoid receptor, and hepatic nuclear factor 4alpha. *Mol. Pharmacol.* **68**: 747-757 (2005).
11) Pitarque, M., Rodriguez-Antona, C., Oscarson, M. and Ingelman-Sundberg, M.: Transcriptional regulation of the human CYP2A6 gene. *J. Pharmacol. Exp. Ther.* **313**: 814-822 (2005).
12) Tirona, R. G., Lee, W., Leake, B. F., Lan, L. B., Cline, C. B., Lamba, V., Parviz, F., Duncan, S. A., Inoue, Y., Gonzalez, F. J., Schuetz, E. G. and Kim, R. B.: The orphan nuclear receptor HNF4alpha determines PXR- and CAR-mediated xenobiotic induction of CYP3A4. *Nat. Med.* **9**: 220-224 (2003).
13) Barbier, O., Girard, H., Inoue, Y., Duez, H., Villeneuve, L., Kamiya, A., Fruchart, J. C., Guillemette, C., Gonzalez, F. J. and Staels, B.: Hepatic expression of the UGT1A9 gene is governed by hepatocyte nuclear factor 4alpha. *Mol. Pharmacol.* **67**: 241-249 (2005).
14) Kitamura, Y., Moriguchi, M., Kaneko, H., Morisaki, H., Morisaki, T., Toyama, K. and Kamatani, N.: Determination of probability distribution of diplotype configuration (diplotype distribution) for each subject from genotypic data using the EM algorithm. *Ann. Hum. Genet.* **66**: 183-193 (2002).
15) Hani, E. H., Suaud, L., Boutin, P., Chevre, J. C., Durand, E., Philippi, A., Demenais, F., Vionnet, N., Furuta, H., Velho, G., Bell, G. I., Laine, B. and Froguel, P.: A missense mutation in hepatocyte nuclear factor-4 alpha, resulting in a reduced transactivation activity, in human late-onset non-insulin-dependent diabetes mellitus. *J. Clin. Invest.* **101**: 521-526 (1998).
16) Furuta, H., Iwasaki, N., Oda, N., Hinokio, Y., Horikawa, Y., Yamagata, K., Yano, N., Sugahiro, J., Ogata, M., Ohgawara, H., Omori, Y., Iwamoto, Y. and Bell G. I.: Organization and partial sequence of the hepatocyte nuclear factor-4 alpha/MODY1 gene and identification of a missense mutation, R127W, in a Japanese family with MODY. *Diabetes* **46**: 1652-1657 (1997).
17) Yang, Q., Yamagata, K., Yamamoto, K., Cao, Y., Miyagawa, J., Fukamizu, A., Hanafusa, T. and Matsuzawa, Y.: R127W-HNF-4alpha is a loss of function mutation but not a rare polymorphism and causes Type II diabetes in a Japanese family with MODY1. *Diabetologia* **43**: 520-524 (2000).
18) Zhu, Q., Yamagata, K., Miura, A., Shihara, N., Horikawa, Y., Takeda, J., Miyagawa, J. and Matsuzawa, Y.: T130I mutation in HNF-4alpha gene is a loss-of-function mutation in hepatocytes and is associated with late-onset Type 2 diabetes mellitus in Japanese subjects. *Diabetologia* **46**: 567-573 (2003).
19) Ek, J., Rose, C. S., Jensen, D. P., Glumer, C., Borch-Johnsen, K., Jorgensen, T., Pedersen, O. and Hansen, T.: The functional Thr130Ile and Val255Met polymorphisms of the hepatocyte nuclear factor-4alpha (HNF4A): gene associations with type 2 diabetes or altered beta-cell function among Danes. *J. Clin. Endocrinol. Metab.* **90**: 3054-3059 (2005).
20) Sakurai, K., Seki, N., Fujii, R., Yagui, K., Tokuyama, Y., Shimada, F., Makino, H., Suzuki, Y., Hashimoto, N., Saito, Y., Egashira, T., Matsui, K. and Kanatsuka, A.: Mutations in the hepatocyte nuclear factor-4alpha gene in Japanese with non-insulin-dependent diabetes: a nucleotide substitution in the polypyrimidine tract of intron 1b. *Horm. Metab. Res.* **32**: 316-320 (2000).
21) Rissanen, J., Wang, H., Miettinen, R., Karkkainen, P., Kekalainen, P., Mykkanen, L., Kuusisto, J., Karhapaa, P., Niskanen, L., Uusitupa, M. and Laakso, M.: Variants in the hepatocyte nuclear factor-1alpha and -4alpha genes in Finnish and Chinese subjects with late-onset type 2 diabetes. *Diabetes Care* **23**: 1533-1538 (2000).
22) Hinds, D. A., Stuve, L. L., Nilsen, G. B., Halperin, E., Eskin, E., Ballinger, D.G., Frazer, K. A. and Cox, D. R.: Whole-genome patterns of common DNA variation in three human populations. *Science* **307**: 1072-1079 (2005).

Drug Metab. Pharmacokinet. 22 (1): *SNP1* (48)-*SNP13* (60) (2007).

SNP Communication

Genetic Variations and Haplotype Structures of the ABC Transporter Gene ABCC1 in a Japanese Population

Hiromi FUKUSHIMA-UESAKA[1], Yoshiro SAITO[1,2,*], Masahiro TOHKIN[1,3], Keiko MAEKAWA[1,2], Ryuichi HASEGAWA[3], Manabu KAWAMOTO[4], Naoyuki KAMATANI[4], Kazuko SUZUKI[5], Tatsuo YANAGAWA[5], Hiroshi KAJIO[6], Nobuaki KUZUYA[6], Kazuki YASUDA[7] and Jun-ichi SAWADA[1,2]

[1]*Project Team for Pharmacogenetics,* [2]*Division of Biochemistry and Immunochemistry,* [3]*Division of Medicinal Safety Science, National Institute of Health Sciences, Tokyo, Japan* [4]*Division of Genomic Medicine, Department of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo, Japan,* [5]*Nerima General Hospital, Tokyo, Japan* [6]*Division of Endocrine and Metabolic Diseases, the Hospital,* [7]*Department of Metabolic Disorder, Research Institute, International Medical Center of Japan, Tokyo, Japan*

Full text of this paper is available at http://www.jstage.jst.go.jp/browse/dmpk

Summary: Multidrug resistance-related protein 1 (MRP1), an ATP-binding cassette transporter encoded by the *ABCC1* gene, is expressed in many tissues, and functions as an efflux transporter for glutathione-, glucuronate- and sulfate-conjugates as well as unconjugated substrates. In this study, the 31 exons and their flanking introns of *ABCC1* were comprehensively screened for genetic variations in 153 Japanese subjects to elucidate the linkage disequilibrium (LD) profiles and haplotype structures of *ABCC1* that is necessary for pharmacogenetic studies of the substrate drugs. Eighty-six genetic variations including 31 novel ones were found: 1 in the 5′-flanking region, 1 in the 5′-untranslated region (UTR), 20 in the coding exons (9 synonymous and 11 nonsynonymous variations), 4 in the 3′-UTR, and 60 in the introns. Of these, eight novel nonsynonymous variations, 726G>T (Trp242Cys), 1199T>C (Ile400Thr), 1967G>C (Ser656Thr), 2530G>A (Gly844Ser), 3490G>A (Val1164Ile), 3550G>A (Glu1184Lys), 3901C>T (Arg1301Cys), and 4502A>G (Asp1501Gly), were detected with an allele frequency of 0.003. Based on the LD profiles, the analyzed regions of the gene were divided into five LD blocks (Blocks -1 and 1 to 4). The multiallelic repeat polymorphism in the 5′-UTR was defined as Block -1. For Blocks 1, 2, 3 and 4, 32, 23, 23 and 13 haplotypes were inferred, and 9, 7, 7 and 6 haplotypes commonly found on ≥10 chromosomes accounted for ≥91% of the inferred haplotypes in each block. Haplotype-tagging single nucleotide polymorphisms for each block were identified to capture the common haplotypes. This study would provide fundamental and useful information for the pharmacogenetic studies of MRP1-dependently effluxed drugs in Japanese.

Key words: ***ABCC1*; genetic variation; amino acid change; haplotype; haplotype tagging SNP**

As of August 8, 2006, the novel variations reported here are not found in the database of Japanese Single Nucleotide Polymorphisms (http://snp.ims.u-tokyo.ac.jp/), dbSNP in the National Center for Biotechnology Information (http://www.ncbi.nlm.nih gov/SNP/), or PharmGKB Database (http://www.pharmgkb.org/).

This study was supported in part by the Program for the Promotion of Fundamental Studies in Health Sciences and the Health and Labor Sciences Research Grants from the Ministry of Health, Labor and Welfare.

Introduction

The multidrug resistance-related protein 1 (MRP1) encoded by the ATP-binding cassette transporter C1 gene (*ABCC1*) belongs to a superfamily of ABC transporters.[1,2] MRP1 was originally identified as the overexpressed transporter cloned from a doxorubicin-selected lung cancer cell line H69AR.[3] The *ABCC1* gene, encod-

Received; August 8, 2006, Accepted; October 5, 2006

**To whom correspondence should be addressed:* Yoshiro SAITO, Ph.D., *Division of Biochemistry and Immunochemistry, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. Tel.: 81-3-5717-3831, Fax.: 81-3-5717-3832, E-mail: yoshiro@nihs.go.jp*

ing a 1531-amino acid, 190 kDa membrane protein, consists of 31 exons and spans over 200 kb at chromosome 16p13.1.[1,4] The protein has 17 transmembrane helices (TM) in three membrane spanning domains (consisting of 5, 6 and 6 TMs) and 2 large cytoplasmic domains between TM11 and TM12, and downstream of TM17.[1,2] These two cytoplasmic domains contain nucleotide binding domains where binding and hydrolysis of ATP occurs to facilitate substrate transport. Two sequence motifs in the nucleotide binding domain are well conserved among all ABC transporters and called Walker A and Walker B. The Walker A motif is involved in the binding of the β-phosphate of ATP, while the Walker B motif interacts with Mg^{2+} ion. In addition, a third conserved motif known as the ABC signature sequence (LSGGQ), located between the Walker A and Walker B motifs, has been identified as a possible binding site for the γ-phosphate of ATP.[2]

MRP1 protein is expressed in many tissues throughout the body at relatively high levels in the heart, adrenal gland, lung, and skeletal muscle, and at medium levels in the liver and kidney.[5] In most tissues, MRP1 is localized to the basolateral membranes of cells for the efflux of its substrates into the blood and thus might play a protective role against toxic substances and metabolites.

In vitro studies have indicated that a number of anticancer drugs are good substrates for MRP1, such as doxorubicin, vincristine and methotrexate. Therefore, MRP1 is thought to be involved in drug resistance in cancer patients.[1,2] In addition, an anti-diabetic drug glibenclamide, but not tolbutamide, is known to inhibit transporting activity of MRP1.[6] MRP1 is also a primarily active transporter of glutathione-, glucuronate- and sulfate-conjugates such as the inflammatory mediator leukotriene C_4 (LTC_4).[1,2] Many unconjugated substrates are also known to be transported concurrently with reduced glutathione. Besides, glutathione disulfide is also transported by MRP1. Thus, MRP1 is thought to be involved in the tissue distribution and elimination of drugs and organic anions, and possibly in redox homeostasis.[1,2]

Genetic polymorphisms in the metabolizing enzymes and transporters are known to influence drug metabolism and disposition. As for the *ABCC1* gene, several polymorphisms with functional significance have been found. In Caucasian populations, 1299G>T (Arg433Ser), 1898G>A (Arg633Gln), 2012G>T (Gly671Val), and 4535C>T (Ser1512Leu) have been reported.[7-9] In addition, Arg433Ser decreases the transport activity for LTC_4 and estrone sulfate, but not for estradiol 17β-glucuronide, *in vitro*.[10] Ito *et al.* found 16 genetic polymorphisms, including 4 nonsynonymous and 8 synonymous ones, in 48 Japanese subjects.[11] An *in vitro* functional study showed that one of the nonsynonymous variations, 2168G>A (Arg723Gln), leads to reduced transport activity for LTC_4, estradiol 17β-glucuronide and methotrexate.[12] However, no haplotype analysis has been reported for the Japanese population. Haplotype is the linked combinations of genetic polymorphisms on the same chromosome and has been shown to sometimes render higher associations with clinical parameters such as drug responses and adverse effects than individual polymorphisms.[13] This information may also be useful for identification of the real functionally-relevant polymorphisms from the linked polymorphisms.

In this study, we searched for genetic variations in *ABCC1* by resequencing all the 31 exons and their surrounding introns of 153 Japanese subjects. The detected variations were then used to perform linkage disequilibrium (LD) and haplotype analyses to identify the haplotype-tagging single nucleotide polymorphisms (htSNPs) that are sufficient to capture common haplotypes in Japanese subjects.

Materials and Methods

Human genomic DNA samples: One hundred fifty-three Japanese subjects participating in this study consisted of 86 diabetic patients administered glimepiride and 67 healthy volunteers. Genomic DNA was extracted from blood leukocytes of the diabetic patients, or from Epstein-Barr virus transformed B lymphocytes derived from the healthy volunteers. The ethical review boards of the International Medical Center of Japan, Nerima General Hospital, Tokyo Women's Medical University and National Institute of Health Sciences approved this study. Written informed consent was obtained from all subjects.

PCR conditions for DNA sequencing: First, two sets of multiplex PCR were performed to amplify all 31 exons of *ABCC1* from 100 ng of genomic DNA using 1.25 units of Z-Taq (Takara Bio Inc., Shiga, Japan) with 0.20 μM each of the mixed primers (Mix 1 and Mix 2) designed in the intronic regions as listed in **Table 1** ("1st PCR"). Mix 1 contained the primers for exons 1, 2 to 5, 8 to 12, 13, 14, and 20 to 23, and Mix 2 for exons 6, 7, 15 to 19, 24 to 26, and 27 to 31. The first PCR conditions were 30 cycles of 98°C for 5 sec, 55°C for 5 sec, and 72°C for 190 sec. Next, each exon, except for exons 1, 4, 22 and 23, was amplified separately in the 2nd PCR using the 1st PCR product as a template by Ex-Taq (0.625 units, Takara Bio Inc.) with the primer sets listed in "2nd PCR" in **Table 1**. Because of high GC contents, exons 1, 4, 22 and 23 were amplified using 0.25 units of LA-Taq (Takara Bio Inc.) with GC buffer II (exon 1) or GC buffer I (exons 4, 22 and 23) using 0.5 μM of the primers listed in **Table 1**. The second PCR conditions were 94°C for 5 min, followed by 30 cycles of 94°C for 30 sec, 55°C for 1 min, and 72°C for

Table 1. Primers used for sequencing *ABCC1*

		Amplified or sequenced region	Forward primer (5′ to 3′)	Reverse primer (5′ to 3′)	Amplified length (bp)
1st PCR	Mix 1	Exon 1	GAGGAGAGGAGAAAAGAAAGGCATC	TTAGTAGAGACGGGGGTCAATCCAT	4,788
		Exon 2 to 5	GTGGAGGTTTATTCTTGGGCAGGTA	CCTCCTCCTCCATACTGATGCCCAC	11,757
		Exon 8 to 12	CTCCTGAGTTCAAGCGATTCTCCTT	CCTATGAGACACTGACAATCCACAC	14,059
		Exon 13 and 14	CTCTTGACATCGGAGCCTGGAAAAT	GGGGATGGGCTGGGAAACGATAAAT	5,445
		Exon 20 to 23	CTTGGTGTCCTTATCAGTGTCCTTC	GGCAGGGGAATCACTTGAACTGGGA	13,579
	Mix 2	Exon 6 and 7	TAAAAAGCACGCCCAGCCTTGAATG	TTTCCTTACAATGCTCAGTCACAGA	6,789
		Exon 15 to 19	GAAGAGTAAAAGAAGACAGAGGTGC	GCTGAACTGCTTTATGAGGGAGGAA	15,848
		Exon 24 to 26	AGCCTTGTTGTTCCCATTTTACAGA	ACTTAGTGAAACCCCCATCTCTACT	5,242
		Exon 27 to 31	ACCTCCACTTGCTCTTTTGAAATAC	CTCTCTGGACTAACATAAAGGGATT	13,210
2nd PCR		Exon 1	CGTTATTTTCCCCTGGTGAC	GGTTGTTTTTTACAGACGGGA	838
		Exon 2	GTGTTTTTCATAGAGGCAGC	ACTAAGCGGCAGAGCAAAGA	800
		Exon 3	GCAGGCTGATTACAGACCAT	AGCCTGGATGAAAGCGAGAT	534
		Exon 4	GAGTAGTTTTGATGTAGTCTATGGC	GCTGGGGCTGATTTAGTGAGGATTT	719
		Exon 5	GAGATGGAGTTTTGCTCTGTCACAC	CAGTATAATACAACAGGGAGTGCCG	876
		Exon 6	TCTGGAAGGAAGAGTGAGCA	AGCACAGGCAGATTCAAGAC	312
		Exon 7	AGAGGGGAGCAGCATCAGCA	TTCTCTGCCCAGTTGCTTTA	445
		Exon 8	AGCCCGTGTTTGTAACCACT	GAGACCAGCCTGACCAACAA	580
		Exon 9	TTTGGGTGACTTGCTGGTGA	TACAATGCCTGATGCGTGCT	605
		Exon 10	TGAGTTCAAGGTTGGGAGGC	CGTCCATTGTAGGGTAGTTA	650
		Exon 11	CAGGAATGAAACCACAGGCT	CGGGGCATCAGCATTTGTTA	419
		Exon 12	GTGAAACCCCATCTCTATTG	ACCTGGGCAACATAGTGACC	462
		Exon 13	CGGTCGTTGATTTATCCAGT	CAAGTGCGAGTTCTCTGTCA	563
		Exon 14	GTGGGACCTTCAGAAATAAG	AGTGTGAGACAGGACAGAAG	428
		Exon 15	CAACCCCATCTTACAAGGAC	TTCTCACGACAGCCTAAGCA	507
		Exon 16	GCCTTCAGTGTTTAGTACAG	TGCTGGGAGACAGATGGAAA	502
		Exon 17	TGGCATCTGTTGTCCTTTGT	AAGTGAGACCTGAGCCACAC	537
		Exon 18	CCTGGTCTCAAGCAGTCCTT	GGACTTACCCCTATAAAGAC	496
		Exon 19	AGACCTGAGTTTTGCCCACC	CTCATCAGGTCCAAGGTCAT	603
		Exon 20	CTCAGTGGATGGAGCCTTCT	AAGAGTGCCACATTCCTTCT	531
		Exon 21	AACACTCCGTCTCTTATGCC	TCAAATCCCAGTTCTGCCGC	464
		Exon 22	CGACTTTGTTTACTGCTGAC	AAAGCACTCAAACACCCACT	566
		Exon 23	GAGACGGAGTTTCACTGTGTTGGC	GACAGGTGGAAAAGTAAAAC	779
		Exon 24	TCATTGGTGGTGCTATTCCT	AGTGCCGCTGTCTGCTCTTA	556
		Exon 25	CTGCGGAGTTACTTGAGTTA	TGTCAAATCCGTCTCCTGCT	422
		Exon 26	TAGTGACTGATGGGGTTTCG	CAGCATCCCACAGTCTCGTA	541
		Exon 27	AGGGGGTCATTTGGGGAATA	TCATTTGGTCTTCAGGCTGT	542
		Exon 28	AGCCGAGTCATTCCTTTTGG	AAAAGAACGATGAAGTAGGG	510
		Exon 29	GTTCAAGTGATTCTCCTGCCTCAGT	CCTGGATTGAGACCCGTTTTACAGA	703
		Exon 30	ACACAGATGTTGGGAGTGGA	AGGGATAAGGACAGTGTTGA	436
		Exon 31	CTGACCCGAAGCAGTGACTT	CGGATGCCAAGGGAGAGAAT	1,429
Sequencing*		Exon 1	AAAGTGGTCGCAGGGTGTGT	GTCACCCAAGTTTCCCCCAT	
		Exon 2	GGAGCCTTGTCTGTTTCTTC	AAGGAACTTAGGGTCAACTA	
		Exon 3	GCAGGCTGATTACAGACCAT	AAAAAAAAAAGGCTACATTT	
		Exon 4	AGCCTGGGTGACAAGAGTGA	GGGCTGATTTAGTGAGGATT	
		Exon 5	GGATTACAGTTGCCCACCAC	GCCAAGTGAGAAACCTACAG	
		Exon 12	AAGTTTATGAGAAAAATAGC	ACCTGGGCAACATAGTGACC	
		Exon 22	GTTTACTGCTGACTTTGTTG	ACTCAAACACCCACTCTACA	
		Exon 23	TTATTATTATTAGAAGTTGGGAGTC	AAAACTTAGGAAAAAACTGC	
		Exon 26	AAAGGAAAGTCAAGTACGCC	CAGCATCCCACAGTCTCGTA	
		Exon 29	TACAGGCGTGAACCACCGTA	ACAGATACAGAAACTGAGGC	
		Exon 31	GACTTGCCCAGGTCAGTTGT	CCCCAAGGAAATGAAGCGTT	
			TCTTTGAGATGCTTCTGGCT	GTGGGAACAGTAATAACAGC	
			GCTGTTATTACTGTTCCCAC	TAACATCTAAAAACAAGGAA	

*Only primers which were not the same as those used for the 2nd PCR were shown.

2 min, and then a final extension for 7 min at 72°C. The PCR products were then treated with a PCR Product Pre-Sequencing Kit (USB Co., Cleveland, OH, USA) and directly sequenced on both strands using an ABI BigDye Terminator Cycle Sequencing Kit ver. 3.1 (Applied Biosystems, Foster City, CA, USA) with the sequencing primers listed in **Table 1** (Sequencing). Excess dye was removed by a DyeEX96 kit (Qiagen, Hilden, Germany) and the eluates were analyzed on an ABI Prism 3730 DNA Analyzer (Applied Biosystems).

All novel SNPs were confirmed by repeated sequencing analyses of PCR products generated by a new genomic DNA amplification. Under the conditions used, up to 400 bases upstream of the translational start codon, all of the exons and their flanking introns were successfully sequenced for all the subjects. The genomic and cDNA sequences of *ABCC1* obtained from GenBank (NT 010393.15 and NM 004996.2, respectively) were used as the reference sequences. The nucleotide positions based on the cDNA sequence were numbered from the adenine of the translational initiation site or the nearest exons.

Linkage disequilibrium (LD) and haplotype analyses: Hardy-Weinberg equilibrium and LD analyses were performed by SNPAlyze software ver. 3.1 (Dynacom Co., Yokohama, Japan), and pairwise LDs between variations were obtained for the |D'| and rho square (r^2) values. Some of the haplotypes were unambiguous from subjects with homozygous variations at all sites or a heterozygous variation at only one site. The diplotype configurations (a combination of haplotypes) were inferred by an expectation-maximization based program LDSUPPORT software, which determines the posterior probability distribution of the diplotype for each subject based on the estimated haplotype frequencies.[14] Haplotypes without any amino acid change were designated as **1*, and the nonsynonymous SNP-bearing haplotypes were numerically numbered. Subtypes were named with small alphabetical letters in the order of their frequencies. The ambiguous haplotypes inferred in only one subjects are grouped as "*others*" (**1* group) in **Tables 3 to 6**. The PHASE program (ver. 2.1) was also used for inferring haplotypes of the block with a microsatellite marker detected.[15,16]

Results and Discussion

***ABCC1* variations found in a Japanese population:** We found 86 genetic variations, including 31 novel ones, from 153 Japanese subjects (**Table 2**). Of them, 1 was located in the 5'-flanking region, 1 in the 5'-untranslated region (UTR), 20 in the coding exons (9 synonymous and 11 nonsynonymous variations), 4 in the 3'-UTR, and 60 in the introns. Since we did not find statistically significant difference in the allelic distributions between diabetic patients and healthy volunteers ($p > 0.05$ by χ^2 test or Fisher's exact test) only except for IVS28-266C>G ($p = 0.027$), we analyzed all the variations as one group. As for IVS28-266C>G, the frequency of the minor allele in diabetic patients (0.267) was somewhat lower than that of the healthy volunteers (0.388). All the observed allele frequencies were in Hardy-Weinberg equilibrium ($p > 0.05$).

Eight novel nonsynonymous variations, 726G>T (Trp242Cys), 1199T>C (Ile400Thr), 1967G>C (Ser656Thr), 2530G>A (Gly844Ser), 3490G>A (Val1164Ile), 3550G>A (Glu1184Lys), 3901C>T (Arg1301Cys), and 4502A>G (Asp1501Gly), were found heterozygously in different subjects with an allele frequency of 0.003. All substituted amino acids were located in the cytoplasmic regions of the MRP1 protein.[1,2,12] Two of the changes occurred in the loop between TM5 and TM6 (Trp242Cys) and between TM7 and TM8 (Ile400Thr). Ser656Thr and Gly844Ser were found 22 residues upstream of the Walker A motif and 52 residues downstream of the Walker B motif in the nucleotide binding domain 1 in the loop between TM11 and TM12, respectively. Both Val1164Ile and Glu1184Lys resided in the loop between TM15 and TM16. Arg1301Cys was located 26 residues upstream of the Walker A motif, while and Asp1501Gly was 47 residues downstream of the Walker B motif, in the nucleotide binding domain 2 in the C-terminal. Using PolyPhen program (http://www.bork.embl-heidelberg.de/PolyPhen) to predict the functional effects of amino acid substitutions, two substitutions, Trp242Cys and Gly844Ser, were expected to alter the protein function based on the PSIC (position specific independent count) profile score differences derived from multiple alignments. The loop between TM5 and TM 6 where Trp242 resides is known to be important for interaction with glutathione.[17] Furthermore, Trp242 is located near the regions important for LTC_4 binding (residues 260-274)[18] and LTC_4 transporting activity of MRP1 proteins (amino acids 223-232).[19] As for Ile400Thr, Lys396 mutation to Glu or Ile, located 4 residues upstream, was shown to cause a reduced transport activity.[20] The functional significance of these 8 novel variations should be clarified in the future studies.

We also detected three known nonsynonymous variations, 218C>T (Thr73Ile), 2168G>A (Arg723Gln), and 3173G>A (Arg1058Gln) at frequencies of 0.007, 0.065 and 0.003, respectively. These frequencies were similar to those found in the earlier reports for Japanese[11] and Chinese.[21] One of the variations, Arg723Gln, leads to reduced transport activities for LTC_4, estradiol 17β-glucuronide and methotrexate.[12] We did not detect three previously reported variations: 2012G>T (Gly671Val; found with approximately 0.03 frequency in Caucasians), 3140G>C (Cys1047Ser; 0.05 in African-Americans), and 4535C>T (Ser1512Leu; 0.03 in Caucasians).[8,9,12] These SNPs might be ethnic-specific.

A known microsatellite marker, GCC repeats in the 5'-UTR, was also detected: 9 repeats (1 chromosome), 10 (1), 11 (49), 12 (20), 13 (154), 14 (52), 15 (12), 16 (11), 17 (3), 18 (1), 21 (1) and 23 (1). This polymorphism was first reported from an Italian group ranging from 7 to 14 repeats, but their transcriptional activities did not change between 7 and 14 repeats in an *in vitro* reporter

Table 2. Summary of *ABCC1* variations detected in this study

SNP ID		Pharm GKB[b]	Refe-rence	Location	Position		Nucleotide change	Amino acid change	Frequency		
This Study	dbSNP (NCBI)				NT_010393.15	From the translational initiation site or from the end of the nearest exon[c]			Total (n = 153)	Diabetic patients (n = 86)	Healthy volunteer (n = 67)
MPJ6_AC1001[a]				5′-flanking	7356447	−241	GAGACGCGCGAGG>ATGAGCGGGCGCC		0.020	0.017	0.022
MPJ6_AC1002-013			22	5′-UTR	7356549_7356569	−139_−119	CCGGCTCCCTGC/(GCC)9-23/AGCGCTAGCGCC			(see the text)	
MPJ6_AC1014[a]				Intron 1	7414381	IVS1-371	TAGGCTGGTCTCG>AAACTCCCAGCCT		0.003	0.006	0.000
MPJ6_AC1015[a]				Intron 1	7414742	IVS1-10	TCGCCTGTGTTTG>TTGTTCGCAGGAC		0.003	0.006	0.000
MPJ6_AC1016[a]				Exon 2	7414766	63	CTGGAATGTCACG>ATGGAATACCAGC	T21T	0.003	0.006	0.000
MPJ6_AC1017			11	Exon 2	7414921	218	CTCTCAACAAAAC>TCAAAACTGTAAG	T73I	0.007	0.000	0.015
MPJ6_AC1018	rs4148335		25	Intron 3	7421139	IVS3-288	ATCCCAGCACCTT>GGGGAGGCCAAGG		0.069	0.081	0.052
MPJ6_AC1019	rs4148336		25	Intron 3	7421231	IVS3-196	TAAAAATACAAAA>CATTAGCTAGGCA		0.069	0.081	0.052
MPJ6_AC1020	rs4148337		25	Intron 3	7421361	IVS3-66	CTCCAGCCTGGGT>CGACAAGAGTGAA		0.333	0.355	0.306
MPJ6_AC1021[a]				Intron 4	7421606	IVS4+42	TTCAGTGGACCCG>AGAGGGAGAGATG		0.003	0.000	0.007
MPJ6_AC1022	rs34085651			Intron 4	7423180	IVS4-252	CAGCCTCCGCCTC>TCTGGGTTCAAGT		0.069	0.081	0.052
MPJ6_AC1023	rs185005	#	25	Intron 4	7423323	IVS4-109	AACTCTTGACCTG>CAGGTGTTCTGCC		0.320	0.297	0.351
MPJ6_AC1024	rs246215	#	25	Intron 4	7423332	IVS4-100	CCTCAGGTGTTCC>TGCCTGTCTCGGC		0.320	0.297	0.351
MPJ6_AC1025[a]				Intron 5	7423677	IVS5+120	CCTCAACCCCTGA>TTACAGGGAATAT		0.013	0.006	0.022
MPJ6_AC1026[a]				Intron 5	7423818	IVS5+261	ACTTGGCATGGTG>AATTTTGGAAAAT		0.003	0.000	0.007
MPJ6_AC1027	rs3837750			Intron 5	7423921_7423922	IVS5+364_+365	GATTAGGCCTAT/delAA/TCCTACAGGGCA		0.333	0.355	0.306
MPJ6_AC1028[a]				Intron 5	7439983	IVS5-62	AAGCTCTGACCTG>AGATGAAAAGTCA		0.003	0.006	0.000
MPJ6_AC1029[b]				Exon 7	7443456	726	CAGTGACCTCTGG>TTCCTTAAACAAG	W242C	0.003	0.006	0.000
MPJ6_AC1030	rs903880	#	9	Intron 7	7443593	IVS7+54	CTCCTTTCCACTC>ACTGTGGCCTCAA		0.059	0.070	0.045
MPJ6_AC1031	rs246232	#	21	Intron 7	7443603	IVS7+64	CTCCTGTGGCCTC>GAATCCAGGATGG		0.418	0.436	0.396
MPJ6_AC1032[a]				Intron 7	7443608	IVS7+69	GTGGCCTCAATCC>TAGGATGGGGCCC		0.003	0.000	0.007
MPJ6_AC1033	rs246221	#	11	Exon 8	7451401	825	GCCGGTGAAGGTT>CGTGTACTCCTCC	V275V	0.366	0.366	0.366
MPJ6_AC1034	rs8187851	#		Exon 9	7452778	1047	TTGCAGGTTGCTC>TATCAAGTTCGTG	L349L	0.003	0.006	0.000
MPJ6_AC1035	rs35587	#	7	Exon 9	7452793	1062	CAAGTTCGTGAAT>CGACACGAAGGCC	N354N	0.366	0.366	0.366
MPJ6_AC1036[a]				Exon 9	7452930	1199	AGACCGCTGTCAT>CTGGGGCTGTCTA	I400T	0.003	0.000	0.007
MPJ6_AC1037	rs35588	#	11	Intron 9	7452957	IVS9+8	GGAAGGTAGGGGA>GCGCTGTGCCATT		0.363	0.360	0.366
MPJ6_AC1038	rs35591		25	Intron 9	7454889	IVS9-189	AGGCACTGAGCAC>GCGCGGATAAGAA		0.363	0.360	0.366
MPJ6_AC1039[a]				Intron 9	7454890	IVS9-188	GGCACTGAGCACC>GGCGGATAAGAAT		0.003	0.006	0.000
MPJ6_AC1040	rs35592		25	Intron 9	7454902	IVS9-176	CGCGGATAAGAAT>CGTGGGCTTTGAG		0.366	0.366	0.366
MPJ6_AC1041	rs4148343		25	Intron 10	7455437	IVS10+198	CAGGTTGATGTTC>ATCCTTGGTGGCA		0.114	0.105	0.127
MPJ6_AC1042[a]				Intron 10	7459543	IVS10-117	TAAAAGTAACACA>GCCTTGCCCTGAA		0.042	0.041	0.045
MPJ6_AC1043	rs3743526	#	25	Intron 11	7459874	IVS11+122	CTGCCAGTTGGAC>GTCACTTGGGGAG		0.114	0.105	0.127
MPJ6_AC1044	rs35595		25	Intron 11	7462950	IVS11-78	TAAAGTTTATGAG>AAAAAATAGCTGG		0.209	0.215	0.201
MPJ6_AC1045	rs3765129	#	11	Intron 11	7462980	IVS11-48	TTGAGTGATGGGC>TTGATCCCAGGGT		0.114	0.105	0.127
MPJ6_AC1046[a]				Intron 12	7463262	IVS12+31	TTCCTGGCCCTCA>CTTGTTTGATGTT		0.007	0.012	0.000
MPJ6_AC1047	rs17265551		23	Intron 12	7463287	IVS12+56	TTATTTTCTCTGC>TGTACCTGAATAT		0.069	0.081	0.052
MPJ6_AC1048	rs4148348	#	23	Intron 12	7475007	IVS12-85	GGGAGGGCCCAAG>ACGCGTCTCCAGG		0.124	0.128	0.119
MPJ6_AC1049	rs35604	#	25	Intron 12	7475055	IVS12-37	TGACTCTCACTCA>GGGGCACAGCAGT		0.245	0.227	0.269
MPJ6_AC1050	rs35605	#	11	Exon 13	7475098	1684	TTGCAGGTGGCCC>TTGTGCACATTTG	L562L	0.245	0.227	0.269
MPJ6_AC1051	rs35606	#	25	Intron 13	7475343	IVS13+105	CACACTCCCGGTC>TGGGCTCCATGAG		0.150	0.134	0.172
MPJ6_AC1052			21	Intron 14	7478780	IVS14+115	ACTCTGCCCAGCC>TGCTTGTCTGCGA		0.020	0.023	0.015
MPJ6_AC1053[a]				Exon 15	7483316	1967	CCTGGGCCAGGAG>CCGACCCTCCCAC	S656T	0.003	0.006	0.000
MPJ6_AC1054[a]				Intron 15	7486189	IVS15-99	TTCTGTCTTTCTC>GTTTCTCTACTTG		0.003	0.000	0.007
MPJ6_AC1055	rs8187863		7	Exon 16	7486300	2001	CATCACCTTCTCC>TATCCCCGAAGGT	S667S	0.003	0.006	0.000
MPJ6_AC1056	rs2301666	#	11	Exon 16	7486306	2007	CTTCTCCATCCCC>TGAAGGTGCTTTG	P669P	0.042	0.047	0.037
MPJ6_AC1057	rs2644983	#		Intron 16	7486595	IVS16+181	TCTGTTCTGTTCC>TGTGCTGTCCTAG		0.278	0.262	0.299

cont.