Authors' Disclosures of Potential Conflicts of Interest

Although all authors completed the disclosure declaration, the following authors or their immediate family members indicated a financial interest. No conflict exists for drugs or devices used in a study if they are not being evaluated as part of the investigation. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Authors	Employment	Leadership	Consultant	Stock	Honoraria	Research Funds	Testimony	Other
Wasaburo Koizumi					Astra Zeneca (A)			
Hiroya Takiuchi					Eli Lilly (A)			
José Baselga			Astra Zeneca (A); Roche (A); Merck KgGA (A)		Merck KgGA (A)	Astra Zeneca (B)		
Dollar Amount Codes (A) < $10,000 (B) $10,000-99,999 (C) ≥ $100,000 (N/R) Not Required								

Author Contributions

Conception and design: Federico Rojo, Josep Tabernero, Joan Albanell, Eric Van Cutsem, Atsushi Ohtsu, Toshihiko Doi, Wasaburo Koizumi, Kuniaki Shirao, Hiroya Takiuchi, José Baselga
Administrative support: Wasaburo Koizumi, José Baselga
Provision of study materials or patients: Federico Rojo, Josep Tabernero, Joan Albanell, Eric Van Cutsem, Atsushi Ohtsu, Toshihiko Doi, Wasaburo Koizumi, Kuniaki Shirao, Hiroya Takiuchi, José Baselga
Collection and assembly of data: Federico Rojo, Josep Tabernero, Joan Albanell, Eric Van Cutsem, Atsushi Ohtsu, Toshihiko Doi, Wasaburo Koizumi, Kuniaki Shirao, Hiroya Takiuchi, José Baselga
Data analysis and interpretation: Federico Rojo, Josep Tabernero, Joan Albanell, José Baselga
Manuscript writing: Federico Rojo, Josep Tabernero, S. Ramon Cajal, José Baselga
Final approval of manuscript: Federico Rojo, Josep Tabernero, Joan Albanell, Eric Van Cutsem, Atsushi Ohtsu, Toshihiko Doi, Wasaburo Koizumi, Kuniaki Shirao, Hiroya Takiuchi, S. Ramon Cajal, José Baselga

GLOSSARY

Akt: Protein kinase B belongs to a pathway that is responsible for cell survival. Following activation (ie, phosphorylation) by PI3K, activated Akt blocks the activity of molecules involved in the apoptotic pathway by in turn phosphorylating them.

EGFR (epidermal growth factor receptor): Also known as HER-1, EGFR belongs to a family of receptors (HER-2, HER-3, HER-4 are other members of the family) and binds to the EGF, TGF-α, and other related proteins, leading to the generation of proliferative and survival signals within the cell. It also belongs to the larger family of tyrosine kinase receptors and is generally overexpressed in several solid tumors of epithelial origin.

Gefitinib: Belonging to the class of tyrosine kinase inhibitors, gefitinib (also known as Iressa) binds to the cytoplasmic region of the EGFR that also binds ATP. By competing with ATP binding that is essential for tyrosine phosphorylation, gefitinib inhibits activation of EGFR and blocks the cascade of reactions leading to cellular proliferation.

Ki67: A marker of proliferation, Ki67 is a protein that is expressed in the nucleus of proliferating cells. Absent only in resting cells, cells in the G1, S, G2, and M phase of the cell cycle express this marker.

$P27^{kip1}$: Belongs to the family of cell cycle regulators, typically known as cyclin-dependent kinase inhibitors (CDKI). Kip1/p27, like other CDKIs, binds cyclin-cdk complexes, leading to cell cycle arrest in the G1 phase of growth.

PI3K: Phosphatidylinositol-3 phosphate kinase (PI3K) adds a phosphate group to PI3, which is a downstream signaling molecule involved in survival/proliferative pathways mediated by growth factors such as the EGF and the PDGFs.

pMAPK (phosphorylated mitogen-activated protein kinase): MAPKs are a family of enzymes that form an integrated network influencing cellular functions such as differentiation, proliferation, and cell death. These cytoplasmic proteins modulate the activities of other intracellular proteins by adding phosphate groups to their serine/threonine amino acids. The phosphorylated form of MAPK is being used as a surrogate to the activated form of the receptor.

Tyrosine kinase inhibitor: Molecules that inhibit the activity of tyrosine kinase receptors. They are small molecules developed to inhibit the binding of ATP to the cytoplasmic region of the receptor (eg, gefitinib), thus further blocking the cascade of reactions that is activated by the pathway.

ERRATA

The September 10, 2006, article by Rojo et al entitled, "Pharmacodynamic Studies of Gefitinib in Tumor Biopsy Specimens From Patients With Advanced Gastric Carcinoma" (J Clin Oncol 24:4309-4316, 2006) contained an error in the spelling of S. Ramon y Cajal. It was originally published as S. Ramon Cajal and should have been S. Ramon y Cajal.

DOI: 10.1200/JCO.2006.11.001

The September 10, 2006, Biology of Neoplasia article by Poulin and DeCaprio, entitled "Is There a Role for SV40 in Human Cancer?" (J Clin Oncol 24:4356-4365, 2006) contained an error in the Authors' Disclosure of Potential Conflicts of Interest section. In addition to "Novartis Pharmaceuticals (B)," "Venable LLP (A)" should have been disclosed for James A. DeCaprio in the Consultant category.

DOI: 10.1200/JCO.2006.11.002

The September 20, 2006, article by Ng et al entitled, "Prospective Study of [^{18}F]Fluorodeoxyglucose Positron Emission Tomography and Computed Tomography and Magnetic Resonance Imaging in Oral Cavity Squamous Cell Carcinoma With Palpably Negative Neck" (J Clin Oncol 24:4371-4376, 2006) contained an error. In Figure 1, the x-axis was labeled "Specificity," while it should have been "1–Specificity."

DOI: 10.1200/JCO.2006.11.003

doi: 10.1111/j.1469-1809.2006.00260.x

Genetic Variations and Haplotype Structures of the *ABCB1* Gene in a Japanese Population: An Expanded Haplotype Block Covering the Distal Promoter Region, and Associated Ethnic Differences

K. Sai[1,2,*], M. Itoda[1], Y. Saito[1,3], K. Kurose[1,4], N. Katori[1,5], N. Kaniwa[1,4], K. Komamura[7,8], T. Kotake[9], H. Morishita[9], H. Tomoike[7], S. Kamakura[7], M. Kitakaze[7], T. Tamura[10], N. Yamamoto[10], H. Kunitoh[10], Y. Yamada[11], Y. Ohe[10], Y. Shimada[11], K. Shirao[11], H. Minami[12], A. Ohtsu[13], T. Yoshida[14], N. Saijo[15], N. Kamatani[16], S. Ozawa[1,6] and J. Sawada[1,3]

[1]*Project Team for Pharmacogenetics,* [2]*Division of Xenobiotic Metabolism and Disposition,* [3]*Biochemistry and Immunochemistry,* [4]*Division of Medicinal Safety Science,* [5]*Division of Drugs,* [6]*Division of Pharmacology, National Institute of Health Sciences, Tokyo, 158-8501*

[7]*Division of Cardiology,* [8]*Department of Cardiovascular Dynamics Research Institute,* [9]*Department of Pharmacy, National Cardiovascular Center, Suita, 565-8565*

[10]*Thoracic Oncology Division,* [11]*Gastrointestinal Oncology Division, National Cancer Center Hospital,* [14]*Genetics Division, National Cancer Center Research Institute, Tokyo, 104-0045*

[12]*Division of Oncology/Hematology,* [13]*Division of GI Oncology/Digestive Endoscopy,* [15]*National Cancer Center Hospital East, Kashiwa, 277-8577 and*

[16]*Division of Genomic Medicine, Department of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo, 162-0054, Japan*

Summary

As functional *ABCB1* haplotypes were recently reported in the promoter region of the gene, we resequenced the *ABCB1* distal promoter region, along with other regions (the enhancer and proximal promoter regions, and all 28 exons), in a total of 533 Japanese subjects. Linkage disequilibrium (LD) analysis based on 92 genetic variations revealed 4 LD blocks with the same make up as previously described (Blocks −1, 1, 2 and 3), except that Block 1 was expanded to include the distal promoter region, and that a new linkage between polymorphisms −1789G>A in the distal promoter region and IVS5 + 123A>G in intron 5 was identified. We re-assigned Block 1 haplotypes, and added novel haplotypes to the other 3 blocks. The reported promoter haplotypes were further classified into several types according to tagging variations within Block 1 coding or intronic regions. Our current data reconfirm the haplotype profiles of the other three blocks, add more detailed information on functionally-important haplotypes in Block 1 and 2 in the Japanese population, and identified differences in haplotype profiles between ethnic groups. Our updated analysis of *ABCB1* haplotype blocks will assist pharmacogenetic and disease-association studies carried out using Asian subjects.

Keywords: *ABCB1*, P-gp, haplotype

Introduction

The *ABCB1* gene, encoding p-glycoprotein (P-gp)/multidrug resistance protein 1 (MDR1), is located on chromosome 7q21-q31 and consists of 28 exons. P-gp (1280 amino acids), a member of the

*Correspondence to: Dr. Kimie Sai, Division of Xenobiotic Metabolism and Disposition, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo, 158-8501, Japan. Tel. +81-3-3700-9478; Fax: +81-3-3707-6950, E-mail: sai@nihs.go.jp

ATP-binding cassette (ABC) transporter superfamily, is a large transmembrane glycoprotein that consists of two transmembrane domains (TMDs) and two nucleotide-binding domains (NBDs). P-gp was initially identified as a component of the multidrug resistance phenotype in cancer cells (Riordan *et al.* 1985), but was later found to be widely expressed in normal epithelial cells of tissues such as the liver, intestine, kidneys, and the blood-brain and testis barriers, as well as in lymphocytes (Fojo *et al.* 1987; Cordon-Cardo *et al.* 1989). It is thought that P-gp plays a role in the protection of these tissues against structurally-unrelated toxic xenobiotics, and can modify the oral bioavailability and renal secretion of a variety of drugs (Hoffmann & Kroemer, 2004). Multiple other physiological functions of P-gp have also been suggested in lipid transport (van Helvoort *et al.* 1996), cholesterol metabolism (Debry *et al.* 1997), inhibition of ceramide-induced apoptosis (Liu *et al.* 2001), and the initiation of immune responses by cytokine release (Drach *et al.* 1996). Moreover, reduced P-gp expression has been linked to cancer (Siegsmund *et al.* 2002) and other diseases such as Parkinson's disease (Furuno *et al.* 2002) and ulcerative colitis (Schwab *et al.* 2003).

With recent advances in genomics research there has been an increasing number of pharmacogenetic studies focused on the *ABCB1* gene. Hoffmeyer *et al.* (2000) showed that a synonymous 3435C>T mutation in exon 26 was associated with reduced P-gp expression in the duodenum, and increased plasma levels of digoxin following its oral administration in healthy volunteers. Thus, the 3435C>T single nucleotide polymorphism (SNP) has become the focus of much attention. However, reports on the role of this common SNP have been very inconsistent, which suggests that other functional polymorphisms may be linked with 3435C>T (Kim, 2002). Further studies revealed that 3435C>T was closely linked to other common polymorphisms, such as 1236C>T (silent) at exon 12 and 2677G>T (Ala893Ser) at exon 21, and that the combinations of these SNPs (i.e. haplotypes) differed greatly between ethnic groups (Kim *et al.* 2001; Kroetz *et al.* 2003; Tang *et al.* 2002, 2004). While an *in vitro* functional study on the nonsynonymous 2677G>T (Ala893Ser) SNP at exon 21 showed that 2677G>T was associated with enhanced P-gp activity (Kim *et al.* 2001), other studies found no association (Kimchi-Sarfaty *et al.* 2002; Morita *et al.* 2003; Kroetz *et al.* 2003). One of these latter studies also revealed that another nonsynonymous SNP, 2677G>A (Ala893Thr), had no impact on P-gp function (Morita *et al.* 2003). Yet several clinical studies have shown that the haplotypes 2677T-3435T and 1236T-2677T-3435T are associated with reduced P-gp activity (Johne *et al.* 2002; Kurata *et al.* 2002; Chowbay *et al.* 2003; Wong *et al.* 2005), and that 2677A-bearing subjects exhibit higher P-gp activity (Yi *et al.* 2004). Studies that found no association between these *ABCB1* SNPs and P-gp expression levels (Goto *et al.* 2002), and other conflicting results, have been summarized in recent review articles (Kim, 2002; Ieiri *et al.* 2004).

Recently, *ABCB1* gene promoter region haplotypes were reported by two Japanese research groups, and revealed the existence of functional haplotypes that resulted in altered P-gp expression (Taniguchi *et al.* 2003; Takane *et al.* 2004). In these studies, haplotypes that included −1789G>A alone or in combination with −145C>G were associated with decreased P-gp expression. However, the reported effects of haplotypes carrying −129T>C and two other linked SNPs on P-gp expression were contradictory, showing reduction and enhancement.

From these findings it is clear that the establishment of detailed *ABCB1* gene haplotype profiles specific for each ethnic group is important. We previously conducted haplotype analysis on 145 Japanese subjects by dividing the *ABCB1* gene into 4 blocks, one of which included the proximal promoter region, and revealed that the *2 haplotype in Block 2, which harbours 1236C>T, 2677G>T and 3435C>T, showed a strong association with reduced renal clearance of irinotecan and its metabolites (Sai *et al.* 2003). However, recent findings on the functional distal *ABCB1* promoter region prompted us to identify the extended haplotypes that encompassed the above promoter region in a larger Japanese population.

In this study, we sequenced the distal *ABCB1* gene promoter regions from 533 Japanese subjects. This region covered approximately 2.5 kb upstream from the translational initiation site, adjacent to the previously described Block 1 region. We found that the promoter region SNPs were closely linked with SNPs located over a relatively wide range (up to intron 5) in Block 1, such

Table 1 Additional primers used for sequencing of the *ABCB1* gene promoter region

Primer name	Forward primer (5′ to 3′)	Primer name	Reverse primer (5′ to 3′)
First amplification[a]			
MDR1-1ZF1	CCTGCTCTGTTTTTCACCGT	MDR1-1ZR1	ATTGGTTTCCTCTATGCAGA
Second amplification			
MDR1-P1F	GAGAGGGACTACTGGTTAGC	MDR1-P1R	TGGTCCATCTGGGGTAAATG
MDR1-P2F	AAGGACTGTTGAAAGTAGCA	MDR1-P2R	TTTGAGACGGAGTCTTGCTT
MDR1-P3F	CAGAGATCATAGGCACAAAT	MDR1-P3R	AAACTTCAGACGTCAGATCA
MDR1-P4F	GAAACATCCTCAGACTATGC	MDR1-P4R	CAGGAGGAATGTTCTGGCTT
Sequencing			
MDR1-P5F	ATTTCTTTGAAGTGCTTGGC	MDR1-P5R	GCCACCACCACTTCTGTCAA
MDR1-P6F	GATCTTTACCTGATGCTCAA	MDR1-P6R	GTGCCTATGATCTCTGTTTT
MDR1-P7F	AGCTCACGCCTGTAATCCCT	MDR1-P1R	TGGTCCATCTGGGGTAAATG
MDR1-P4F	GAAACATCCTCAGACTATGC	MDR1-P8R	AGGAAAAGTACGTGCAATCT
MDR1-P9F	ACGTACTTTTCCTCAGTTTG	MDR1-P9R	ACACGTCTTTCAAAGTTCAC

Other primer sets used were as previously reported (Sai *et al.* 2003).
[a]The same set as previously used for the enhancer and promoter regions.

that it was necessary to re-evaluate the functional significance of Block 1 haplotypes. We also sequenced the same regions as covered by the previous study, including the enhancer region (Geick *et al.* 2001) and all exons and surrounding introns, for an additional 388 subjects. These results allowed us to add novel haplotypes to three other blocks. Lastly, we performed a network analysis on the haplotypes obtained in each block and compared the profile of *ABCB1* haplotypes in Japanese with those of other ethnic groups (Kroetz *et al.* 2003; Takane *et al.* 2004).

Materials and Methods

DNA Samples

All 533 Japanese subjects were patients with either ventricular tachycardia (121 subjects) who were administered an anti-arrhythmic drug (amiodarone) and/or β-blockers, or with various cancers (412 subjects) who were administered an anti-cancer drug (paclitaxel or irinotecan). Genomic DNA was extracted directly from blood leukocytes. This study was approved by the ethical review boards of the National Cardiovascular Center, the National Cancer Center, and the National Institute of Health Sciences. Written informed consent was obtained from all subjects.

DNA Sequencing

Amplification and sequencing of the *ABCB1* gene were performed as previously described (Sai *et al.* 2003), except that the region sequenced included the promoter region up to 2.5 kb upstream from the translational initiation site. For the promoter region, PCR amplification was first performed using the previous primer set that covered from 7 kb upstream of the transcription site to exon 3, and then new primer sets were used for the second PCR and sequencing (Table 1). Amplification and sequencing primers for the other regions and the PCR conditions used were the same as previously reported (Sai *et al.* 2003). Genbank NT_007993.14 was used as the reference sequence. Nucleotide positions were based on cDNA sequence as previously described, with the adenine of the translational initiation site at exon 2 numbered as +1. For 5′-flanking variations intron 1 was skipped for numbering nucleotide positions.

Haplotype and Network Analyses

Linkage disequilibrium (LD) analysis was performed using SNPAlyze software (Dynacom Co., Yokohama, Japan). According to the LD pattern we divided the *ABCB1* gene into 4 blocks following the previously described block partitioning, except for a changed border between Block 1 and Block 2 (IVS5 + 123A>G was shifted from Block 2 to Block 1). Diplotype configurations (combinations of haplotypes) in each block were inferred by LDSUPPORT software, which determined the posterior probability distribution of diplotype configurations for each subject based on estimated haplotype frequencies (Kitamura *et al.* 2002). As Block 1 was expanded we re-defined the Block 1 haplotypes.

For Block 2 haplotypes the previously defined *8c was deleted due to a shift of IVS5 + 123A>G to Block 1. For the rest of the haplotypes we followed the haplotype nomenclature used in our previous study (Sai *et al.* 2003) and added the newly-identified haplotypes consecutively. In our nomenclature the group of haplotypes without amino acid changes or marker SNPs in Block 2 (1236C>T, 2677G>T/A and 3435>T) was defined as *1, and haplotype groups bearing nonsynonymous SNPs or marker SNPs in Block 2 were consecutively numbered as described previously (Sai *et al.* 2003). Novel haplotypes within each haplotype group were designated in descending order of frequency. Haplotypes inferred in only one patient, or ambiguously defined, were described with "?", and some rare variations described as "Others" in Figures 3-5. To allow comparison with previous reports (Taniguchi *et al.* 2003; Takane *et al.* 2004) an additional classification for Block 1 haplotypes was given in Fig. 7, based on marker SNPs of the promoter region (−1789G>A, −1461_−1457delCATCC, −371A>G, −145C>G and −129T>C).

Network analysis of haplotypes was performed to obtain cladograms using Network 4.1.0.9 (www.fluxus-engineering.com.). Network calculations were based on algorithms of the reduced median network (for Blocks −1, 1 and 3) or the median joining network (for Block 2). Haplotypes inferred in only one patient were omitted from the network analysis due to their low predictability.

Results

Additional Genetic Variations

In this study we sequenced the distal promoter region covering approximately 2.5 kb upstream of the translational initiation site in exon 2 in 533 Japanese subjects. We also re-sequenced the enhancer region, and all 28 exons and surrounding regions (the same regions that were sequenced in the previous paper), in an additional 388 subjects. A total of 92 genetic variations were detected in the entire region sequenced in this study. All of the allelic frequencies were in Hardy-Weinberg equilibrium. Since we did not find any apparent differences in SNP frequencies between the two disease types ($P \geq 0.2233$; Fisher's exact test), the data from all subjects were analyzed as one group.

In addition to the variations reported in our previous study we detected 44 further variations, including 35 novel variations, as listed in Table 2. Novel variations included 8 nonsynonymous substitutions: 49T>C(F17L), 144G>T(K48N), 304G>C(G102R), 1342G>A(E448K), 1804G>A(D602N), 2359C>T (R787W), 2719G>A(V907I) and 3043A>G(T1015A); and 2 synonymous substitutions: 354C>T (Y118Y) and 447A>G(K149K); with frequencies ranging from 0.001 to 0.005. Other novel variations in the 5′-flanking region were 11 nucleotide substitutions and one deletion, while in the intronic regions there were 11 nucleotide substitutions, one deletion, and one insertion (Table 2).

The highly polymorphic variations 1236C>T, 2677G>T, 2677G>A, and 3435C>T were detected at frequencies of 0.572, 0.410, 0.183, and 0.440, respectively, which was consistent with our previous observations (Sai *et al.* 2003). In the newly-sequenced promoter region the reported polymorphic variations −1847T>C, −1789G>A, −1461·−1457del CATCC, and −1347T>C were found at frequencies of 0.084, 0.204, 0.030, and 0.084, respectively, which were comparable with frequencies in Japanese in previous reports (Taniguchi *et al.* 2003; Takane *et al.* 2004).

LD analysis was performed using the 92 detected genetic variations, and pairwise rho square (r^2) values for the representative 46 polymorphisms (alleles detected in 5 or more chromosomes), and the results are shown in Fig. 1. With the additional distal promoter region sequence close linkage relationships were observed between −1847T>C, −1347T>C, −371A>G, −129T>C, IVS3 + 36C>T and IVS5 + 76T>G. A close linkage was also detected between −1789G>A in the promoter region and IVS5 + 123A>G in intron 5 (formerly classified as Block 2). Based on these linkage relationships we changed the previous border between Block 1 and Block 2, such that IVS5 + 123A>G was now classified as part of Block 1. The other linkage profiles were the same as previously described, confirming the previous partitioning between Blocks 2 and 3. Similarly, the enhancer region at around 7 kb

Table 2 Additional *ABCB1* variations detected in Japanese

Block	SNP ID		Site	Position		Nucleotide change	Amino acid change	Frequency
	This study[a]	Reference		NT_007933.14	cDNA-based			
Block −1	MDJ6_AB1078	(novel)	5′-Flanking	12472468_12472461	−8128_−8121	GTAAAGTCAGATCTAACCAA/-CTGTTCATTGGT		0.002
	MDJ6_AB1079	(novel)	5′-Flanking	12466729	−2389	CTCCCATAGATAC/TATATAGAACAGA		0.001
	MDJ6_AB1080	b)	5′-Flanking	12466680	−2340	ATGTTGTCAGAGT/CATAGACAAGTTG		0.001
	MDJ6_AB1081	(novel)	5′-Flanking	12466659	−2319	GTTGGTGAAATGG/TCTACATGAGAGC		0.001
	MDJ6_AB1072	b,c)	5′-Flanking	12466187	−1847	GTTTAGGGAGGGT/CTTAAGGCCATTC		0.084
	MDJ6_AB1073	rs12720464[d]	5′-Flanking	12466129	−1789	AATGAAAGGTGAG/AATAAAGCAACAA		0.204
	MDJ6_AB1082	(novel)	5′-Flanking	12466065	−1725	AAGATTAAAAACG/ACATGTAATGAAG		0.001
	MDJ6_AB1083	(novel)	5′-Flanking	12465983	−1643	CAGTGAACAATGC/TTGTACACTTGCA		0.001
	MDJ6_AB1084	(novel)	5′-Flanking	12465806	−1466	GGTCAGGAGATCA/GAGACCATCCTGG		0.002
	MDJ6_AB1085	c)	5′-Flanking	12465801_12465797	−1461_−1457	GGAGATCAAGACCATCC /-TGGCTAACACAG		0.030
Block 1	MDJ6_AB1074	b,c)	5′-Flanking	12465687	−1347	GCAGGAGAATGGT/CGTGAACCCGGGA		0.084
	MDJ6_AB1086	(novel)	5′-Flanking	12465619	−1279	CCTGGGCGACAAA/GGCAAGACTCCGT		0.004
	MDJ6_AB1075	b,c)	5′-Flanking	12465494	−1154	AGAAAAAAATTAT/CGGCTTTGAAGTA		0.001
	MDJ6_AB1087	(novel)	5′-Flanking	12465444	−1104	ATCCTCAGACTAT/CGCAGTAAAAAAC		0.001
	MDJ6_AB1088	(novel)	5′-Flanking	12465421	−1081	ACAAAGTGATTTT/CCTTCTTCTAAAC		0.002
	MDJ6_AB1089	(novel)	5′-Flanking	12465405	−1065	CTTCTAAACTTAT/CGCAATAAACTGA		0.001
	MDJ6_AB1090	(novel)	5′-Flanking	12465326	−986	TCCTCTATGTTCA/GTAAGAAGTAAGA		0.001
	MDJ6_AB1091	(novel)	5′-Flanking	12464967	−627	TTATCATCAAATA/GAAGGATGAACAG		0.002
	MDJ6_AB1092	(novel)	Exon 2	12463728	49	AAGAAGAACTTTT/CTTAAACTGAACA	F17L	0.001
	MDJ6_AB1093	(novel)	Exon 4	12449246	144	TTGGCTTGACAAG/TTTGTATATGGTG	K48N	0.001
	MDJ6_AB1094	(novel)	Exon 5	12433798	304	ATCAATGATACAG/CGGTTCTTCATGA	G102R	0.005
	MDJ6_AB1095	(novel)	Exon 6	12430553	354	TGCCTATTATTAC/TAGTGGAATTGGT	Y118Y	0.001
	MDJ6_AB1096	(novel)	Exon 6	12430460	447	CAAAATTAGAAAA/GCAGTTTTTTCAT	K149K	0.002
	MDJ6_AB1097	(novel)	Exon 12	12413771	1342	TATGACCCCACAG/AAGGGGATGGTGA	E448K	0.001

Table 2 Continued.

Block	SNP ID: This study[a]	SNP ID: Reference	Site	Position: NT_007933.14	Position: cDNA-based	Nucleotide change	Amino acid change	Frequency
	MPJ6_AB1052	e)	Intron 12	12413746	IVS12 +17	GATGACCCATGC**G/A**AGCTAGACCCTG		0.006
	MPJ6_AB1098	(novel)	Intron 12	12413720	IVS12 +43	GGTGATCAGCAG**T/G**CACATTGCACAT		0.001
	MPJ6_AB1099	(novel)	Intron 13	12413353	IVS13 +90	CTACTATAAATC**G/A**GAAGAAGGGAAA		0.001
	MPJ6_AB1100	(novel)	Exon 15	12409538	1804	ATCGCTGGTTTC**G/A**ATGATGGAGTCA	D602N	0.002
	MPJ6_AB1101	(novel)	Intron 15	12408686	IVS15 −95	GTTACTAAACAA**A/G**TTGCTGTTTTCC		0.001
	MPJ6_AB1065	(novel)	Intron 16	12408363	IVS16 +52	CTGTGGTTCCCTA**C/G**TTTGGTGGGCT		0.003
	MPJ6_AB1102	(novel)	Intron 16	12407939	IVS16 −72	TCCTTTACTAAT**T/A**TTTGTGCGTATG		0.001
	MPJ6_AB1103	(novel)	Intron 18	12404862	IVS18 +87	AGTGTAATTGGC**C/T**TTTTAGTAGAAC		0.001
Block 2	MPJ6_AB1104	(novel)	Exon 19	12402898	2359	ATCCTCACCAAG**C/T**GGCTCCGATACA	R787W	0.001
	MPJ6_AB1105	(novel)	Intron 19	12400221	IVS19 −88	GGGGTATAAGTAT**T/C**AAACAAAACTGA		0.001
	MPJ6_AB1106	(novel)	Intron 20	12395242	IVS20 −153	TTCTTACTGTAG**A/G**AACTCAATAAAC		0.001
	MPJ6_AB1107	(novel)	Intron 20	12395172	IVS20 −83	GAATATAGTCTC**A/G**TGAAGGTGAGTT		0.001
	MPJ6_AB1108	(novel)	Intron 21	12384544_12384541	IVS21−73_−76	TTATTTTCATTA**GTCT/**-GTTTTATAGAAT		0.003
	MPJ6_AB1067	(novel)	Exon 22	12384435	2719	AACTTCCGAACC**G/A**TTGTTTCTTTGA	V907I	0.002
	MPJ6_AB1109	(novel)	Intron 22	12384359	IVS22 +9	ACAGGTAATAACC**C/T**GCTGAAGAGTGG		0.001
	MPJ6_AB1076	f)	Exon 24	12380229	2956	GTCTTTGGTGCC**A/G**TGGCCGTGGGGC	M986V	0.001
	MPJ6_AB1110	(novel)	Exon 24	12380142	3043	ATCATTGAAAAAA**A/G**CCCCTTTGATTG	T1015A	0.001
	MPJ6_AB1111	(novel)	Intron 26	12372831_12372834	IVS26 +33_36	ACAGCCTGGGAG-/**CATG**TGGCAGCCTCTC		0.001
Block 3	MPJ6_AB1112	(novel)	Intron 26	12369713	IVS26 −78	ATATAGAATCGTC**C/G**TATCCTACTTTC		0.001
	MPJ6_AB1077	rs2235051[d]	Exon 28	12367931	3747	GTTTCAGAATGGC**C/G**AGAGTCAAGGAG	G1249G	0.002

All *ABCB1* genetic variations in the above list and detected in the previous study (Sai *et al.* 2003) were used for the haplotype analysis in this study.

[a]SNP ID assigned by our project team (MPJ-6).

[b]Taniguchi *et al.* 2003.

[c]Takane *et al.* 2004.

[d]NCBI dbSNP

[e]Itoda *et al.* 2002.

[f]Tanabe *et al.* 2001.

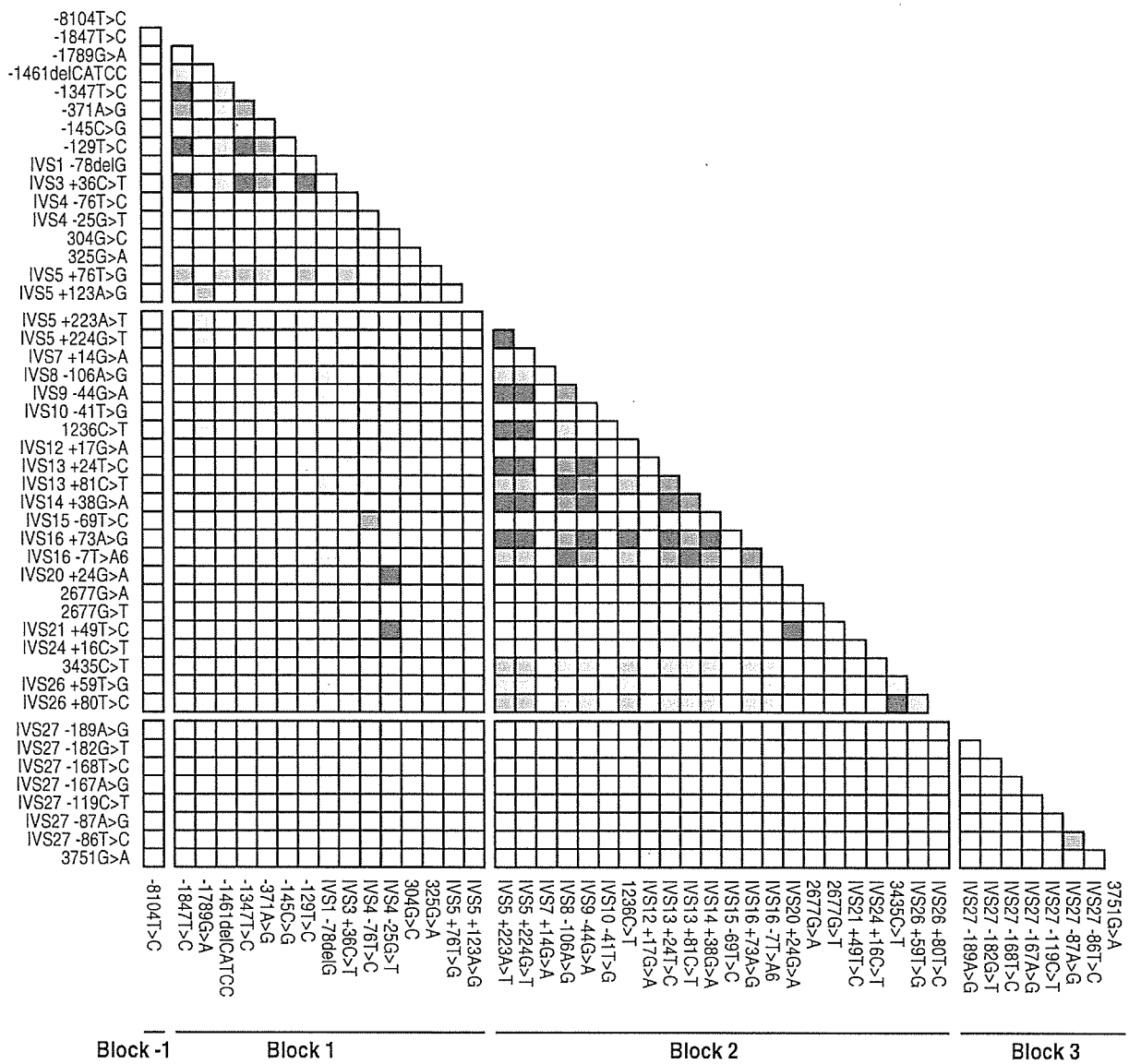

Figure 1 Linkage disequilibrium (LD) analysis of the *ABCB1* gene. Pairwise LD (r^2 values) of the polymorphisms detected in 5 or more chromosomes is shown as a 10-graded blue colour.

upstream of the transcriptional start site was assigned as Block −1 as described previously.

Haplotype Analysis

We estimated the diplotype configurations (haplotype combinations) of all 4 blocks using LDSUPPORT software. Diplotype configurations were obtained at probabilities over 0.9 for 100%, 92%, 95%, and 98% of the subjects for Blocks −1, 1, 2, and 3, respectively.

In Block −1, one novel haplotype, *1d, was identified. Thus Block 1 contained four *1 haplotypes (Fig. 2). The most common haplotype was *1a with a frequency of 0.988.

For Block 1 five haplotype groups consisting of 39 haplotypes were newly assigned. Of the 35 haplotypes in the *1 group, 10 haplotypes were ambiguous and were included as "Others" in Fig. 3. Haplotype groups *2 to *5 were defined by the nonsynonymous SNPs 325G>A(E109K) (*2), 304G>C(G102R) (*3),

49T>C(F17L) (*4) and 144G>T(K48N) (*5). The most frequent haplotype was *1a at a frequency of 0.541, followed by *1b (−1789G>A and IVS5 + 123A>G), *1c (IVS1 − 78delG), and *1d (IVS4 − 25G>T) at frequencies of 0.098, 0.079, and 0.041, respectively. The nonsynonymous *2 and *3 groups occurred at frequencies of 0.017 and 0.005, respectively.

Site		5'-Flanking				
Position		-8128 _8121	-8104	-7970		
Nucleotide change		del CTAA CCAA	T>C	C>T		
Amino acid change					N	Frequency
*1	*1a				1053	0.988
	*1b				2	0.002
	*1c				9	0.008
	*1d				2	0.002

Figure 2 *ABCB1* haplotypes in Block − 1 for 533 Japanese subjects. The haplotype nomenclature followed the definitions used in our previous study (Sai *et al.* 2003). Newly identified haplotypes were consecutively named as shown in boldface.
N: number of chromosomes analyzed.

In Block 2 15 haplotype groups consisting of 61 haplotypes were inferred, including 38 newly-defined haplotypes. Of the 61 haplotypes 24 were detected in only one patient or ambiguously inferred. Ambiguous haplotypes within each group (groups *1, *8, and *10) were indicated as "Others" or "?" in Fig. 4. The most frequent haplotype was the *2 group at a frequency of 0.386, which harboured 1236C>T (exon 12), 2677G>T(A893S) (exon 21) and 3435C>T (exon 26). Groups *1, *10 [2677G>A(A893T)] and *8 (1236C>T) were found at frequencies of 0.216, 0.174 and 0.141, respectively. Other minor haplotype groups were *6 (3435C>T), *9 [1236C>T and 2677G>T(A893S)], *4 (1236C>T and 3435C>T) and *11 [1236C>T and 2677G>A(A893S)] at frequencies of 0.034, 0.020, 0.016, and 0.005, respectively. All these frequencies were comparable with our previous findings (Sai *et al.* 2003). Novel haplotype groups bearing amino acid substitutions were assigned as *12 [1804G>A (D602N)], *13 [2719G>A (V907I)], *14 [1342G>A (E448K)], *15 [2956A>G (M986V)], *16 [3043A>G (T1015A)], and *17 [2359C>T(R787W)],

Site														Ex. 1(5'-UTR)		Int. 1	Ex. 2	Int. 3	Ex. 4	Int. 4		Ex. 5		Int. 5			
Position		-2340	-2319	-1847	-1789	-1466	-1461 _-1457	-1347	-1279	-1081	-1065	-986	-371	-145	-129	IVS1 -78	49	IVS3 +36	144	IVS4 -76	IVS4 -25	304	325	IVS5 +76	IVS5 +123		
Nucleotide change		T>C	G>T	T>C	G>A	A>G	delCA TCC	T>C	A>G	T>C	T>C	A>G	A>G	C>G	T>C	del G	T>C	C>T	G>T	T>C	G>T	G>C	G>A	T>G	A>G		
Amino acid change																	F17L		K48N			G102R	E109K			N	Frequency
*1	*1a																										0.541
	*1b																										0.098
	*1c																										0.079
	*1d																										0.041
	*1e																										0.029
	*1f																										0.028
	*1g																										0.028
	*1h																										0.022
	*1i																										0.017
	*1j																										0.016
	*1k																										0.015
	*1L																										0.012
	*1m																										0.010
	*1n																										0.010
	*1o																										0.004
	*1p																										0.003
	*1q																										0.002
	*1r																										0.002
	*1s																										0.002
	*1t																										0.002
	*1u																										0.002
	*1v																										0.001
	*1w																										0.001
	*1x																										0.001
	*1y																										0.001
	Others																										0.009
*2	*2a																						2				0.017
*3	*3a																					3					0.005
*4	*4a?																4										0.001
*5	*5a?																		5								0.001

Figure 3 *ABCB1* haplotypes in Block 1 for 533 Japanese subjects. Block 1 haplotypes were newly defined due to the change of the Block 1 border. Rare and ambiguous haplotypes (n = 1) are shown with "?". Haplotypes assigned in only one patient or ambiguously inferred are shown as "Others". Sites for nonsynonymous substitutions are indicated by their group-name numbers.
N: number of chromosomes analyzed.

Site		Int. 5		Ex. 6	Int. 6	Int. 7	Int. 8	Int. 9	Int. 10	Ex. 12		Int. 12	Int. 13		Int. 14	Ex. 15	Int. 15		Int. 16		Int. 18		Ex. 19	Int. 19	Int. 20		Ex. 21		Int. 21	Ex. 22	Ex. 24		Int. 24	Ex. 28			
Position		IVS5 +223	IVS5 +224	447	IVS6 -109	IVS7 +14	IVS8 -106	IVS9 -44	IVS10 -41	1236	1342	IVS12 +17	IVS13 +24	IVS13 +81	IVS14 +38	1804	IVS15 -95	IVS15 -69	IVS16 +52	IVS16 +73	IVS16 -76	IVS18 +87	IVS18 -35	2359	IVS19 -88	IVS20 +24	IVS20 -153	2677	2677	IVS21 +49	IVS21 -73_-76	2719	2956	3043	IVS24 +16	3435	IVS26 +59	IVS26 +80	N	Frequency
Nucleotide change		A>T	G>T	A>G	G>T	G>A	A>G	G>A	T>G	C>T	G>A	G>A	T>C	C>T	G>A	G>A	A>G	T>C	A>C	A>G	T>A	C>T	G>C	C>T	T>C	G>A	G>A	G>T	T>C	del GTCT	G>A	A>G	A>G	C>T	C>T	T>G	T>C		
Amino acid change				K149K						G412G	E448K					D602H							R787W				A893T	A893S			V907I	M986V	T1015A		I1145I				
*1	*1a																																					117	0.110
	*1f																																					47	0.044
	*1g																																					27	0.025
	*1L																																					10	0.009
	*1h																																					6	0.006
	*1k																																					6	0.006
	*1m																																					3	0.003
	*1n																																					3	0.003
	*1i																																					2	0.002
	*1o																																					2	0.002
	*1p																																					2	0.002
	Others																																					5	0.005
*2	*2d									2																	2								2				392	0.368
	*2e									2																	2								2				7	0.007
	*2f									2																	2								2				4	0.004
	*2g									2																	2								2				3	0.003
	*2h									2																	2								2				3	0.003
	*2i									2																	2								2				1	0.001
	*2j									2																	2								2				1	0.001
	*2k									2																	2								2				1	0.001
*4	*4a									4																	4								4				16	0.015
	*4c?									4																	4								4				1	0.001
*6	*6a																																					20	0.020
	*6b																																					15	0.013
	*6d?																																					1	0.001
*8	*8a									8																												98	0.092
	*8b									8																												17	0.016
	*8e									8																												10	0.009
	*8h									8																												10	0.009
	*8d									8																												3	0.003
	*8i									8																												3	0.003
	*8k									8																												2	0.002
	*8l									8																												1	0.001
	Others									8																												6	0.006
*9	*9b									9																	9												14	0.013
	*9a									9																	9												7	0.007
*10	*10a																										10												180	0.169
	*10e																										10												1	0.001
	Others																																						5	0.005
*11	*11a									11																	11												4	0.004
	*11c?									11																	11												1	0.001
*12	*12a															12																								2	0.002
*13	*13a																														13								2	0.002
*14	*14a?										14																													1	0.001
*15	*15a?																															15								1	0.001
*16	*16a?																																16							1	0.001
*17	*17a?																							17																1	0.001
*18	*18a?																											18								18				1	0.001

Figure 4 *ABCB1* haplotypes in Block 2 for 533 Japanese subjects. Haplotype nomenclature followed the definitions used in our previous study (Sai *et al.* 2003) except for deletion of *8c. Newly identified haplotypes were consecutively named as shown in boldface. Rare and ambiguous haplotypes (n = 1) are shown with "?". Haplotypes assigned in only one patient or ambiguously inferred are shown as "Others". Marker sites for the *2 to *18 groups are indicated by their group-name numbers.
N: number of chromosomes analyzed.

with frequencies that ranged from 0.002 for *12 and *13, to 0.001 for *14 to *17. Another new haplotype was defined as *18 based on the simultaneous presence of 2677G>A (A893T) and 3435C>T, with a frequency of 0.001. It was also noted that *1f in Block 2 was completely linked with *1d (IVS4 − 25G>T) in Block 1.

In Block 3 three haplotype groups consisting of 21 haplotypes were inferred, including four new haplotypes. Of the 21 haplotypes three were ambiguously inferred and included in "Others" in Fig. 5. The most frequent haplotype was *1a with a frequency of 0.753, followed by *1b (0.176). As observed previously, the rare haplotype groups *2 [3751G>A (V1251I)] and *3 [3587T>G (I1196S)] were observed at frequencies of 0.014 and 0.001, respectively.

We also analyzed the diplotype combinations for all 4 blocks (i.e. the whole gene) for all 533 subjects. The combination patterns were highly diverse with a total of 353 diplotype combinations observed. The frequencies for the majority of diplotypes were less than 0.01. The 10 major combinations are listed in Table 3; all combinations were made up of the major haplotypes in each block.

Network Analysis and Nucleotide Diversity

We performed a network analysis of the haplotypes in each block to obtain cladograms based on the sites and numbers of mutational events. For Block −1 the rare haplotypes, *1b to *1d, appeared to be derived from the major haplotype *1a (Fig. 6a). For Block 1 most of the minor haplotypes were connected to one of the major haplotypes *1a or *1b. However, *1e, *1g, *1h, and *1p were shown to be distant from the above haplotypes (Fig. 6b). Haplotype groups including *1b and the closely related haplotypes *1f, *1i, *1k, *1j, *1m, *1L, and *1q were characterized by the presence of −1789G>A. Of these haplotypes the *1k and *1j subgroups were characterized by the additional SNP −371A>G, while the subgroups *1m,*1L and *1q contained the SNP −145C>G. The separate subgroup that consisted of *1e, *1g, *1h and *1p contained

Site		Int. 26	Ex. 27	Int. 27										Ex. 28			
Position		IVS26 -78	3587	IVS27 +63	IVS27 -189	IVS27 -182	IVS27 -172	IVS27 -168	IVS27 -167	IVS27 -119	IVS27 -87	IVS27 -86	IVS27 -80	3747	3751		
Nucleotide change		C>G	T>G	C>G	A>G	G>T	G>A	T>C	A>G	C>T	A>G	T>C	ins C	C>G	G>A		
Amino acid change			I1196S											G1249G	V1251I	N	Frequency
*1	*1a															803	0.753
	*1b															188	0.176
	*1c															27	0.025
	*1d															5	0.005
	*1f															5	0.005
	*1h															4	0.004
	*1j															4	0.004
	*1e															3	0.003
	*1L															2	0.002
	*1g															1	0.001
	*1i															1	0.001
	*1k															1	0.001
	*1m															1	0.001
	*1o															1	0.001
	*1r															1	0.001
	Others															3	0.003
*2	*2a														2	9	0.008
	*2b														2	6	0.006
*3	*3a		3													1	0.001

Figure 5 *ABCB1* haplotypes in Block 3 for 533 Japanese subjects. Haplotype nomenclature followed the definitions used in our previous study (Sai *et al.* 2003). Newly identified haplotypes were consecutively named as shown in boldface. Haplotypes assigned in only one patient or ambiguously inferred are shown as "Others". Sites for nonsynonymous substitutions are indicated by their group-name numbers.
N: number of chromosomes analyzed.

Table 3 Diversity of block diplotype combinations across the 4 blocks

Block −1	Block 1	Block 2	Block 3	Number of subjects
*1a/*1a	*1a/*1a	*2d/*2d	*1a/*1a	24
*1a/*1a	*1a/*1a	*2d/*2d	*1b/*1a	13
*1a/*1a	*1c/*1a	*10a/*2d	*1a/*1a	12
*1a/*1a	*1a/*1a	*2d/*1e	*1b/*1a	11
*1a/*1a	*1e/*1a	*10a/*2d	*1a/*1a	9
*1a/*1a	*1a/*1a	*2d/*2d	*1c/*1a	9
*1a/*1a	*1b/*1a	*8a/*2d	*1a/*1a	9
*1a/*1a	*1c/*1a	*10a/*1e	*1b/*1a	6
*1a/*1a	*1g/*1a	*10a/*2d	*1a/*1a	6
*1a/*1a	*1d/*1c	*10a/*1f	*1a/*1a	5
*1a/*1a	*2a/*1a	*10a/*2d	*1a/*1a	5
*1a/*1a	*1c/*1a	*10a/*2d	*1b/*1a	5

A total of 353 diplotype-combinations across the 4 blocks were detected in 533 subjects.
The number of subjects for the other combinations was less than 5.

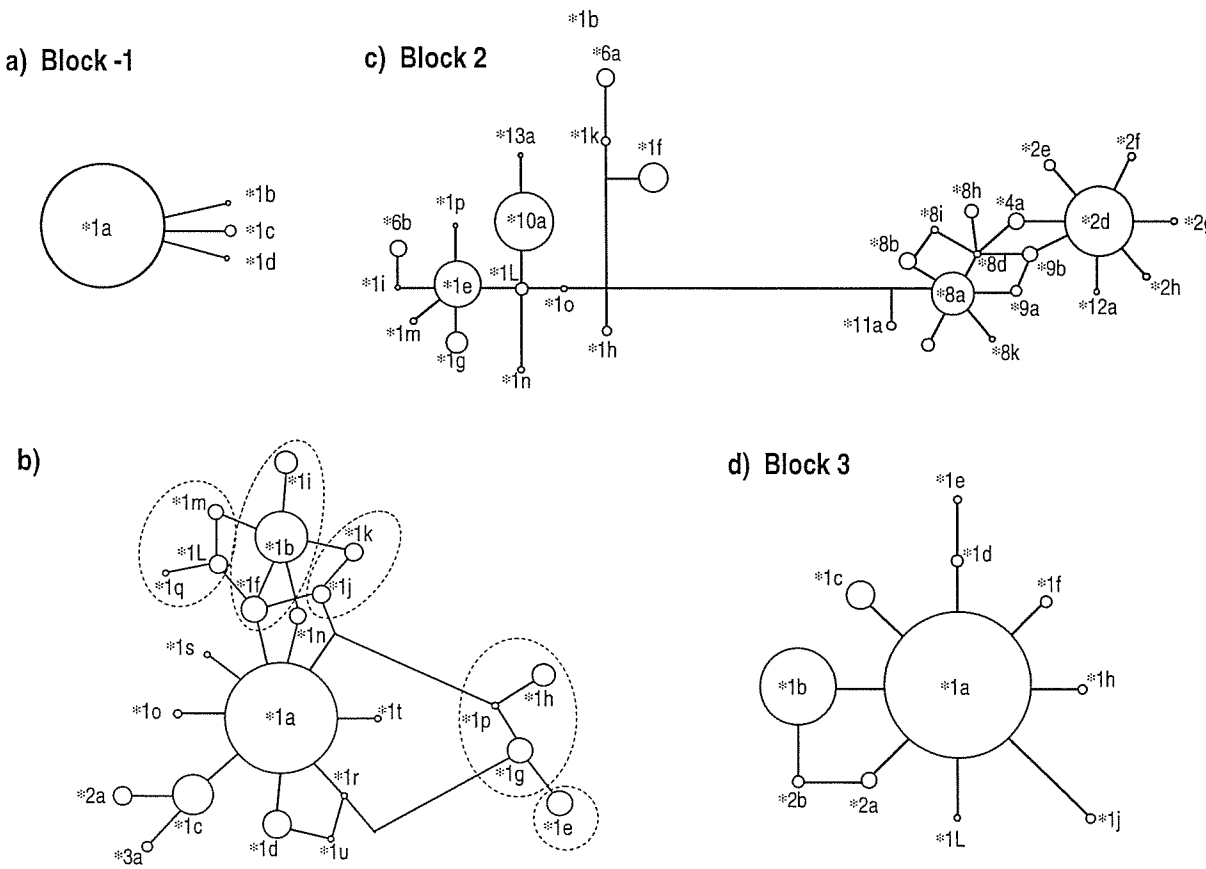

Figure 6 Network analysis of *ABCB1* haplotypes of Block −1 (a), Block 1 (b). Block 2 (c) and Block 3 (d). For each block, the circle area represents the approximate haplotype frequency, and line length between the circles is proportional to the number of mutations. Haplotypes inferred in only one chromosome were omitted from this analysis. The classification by Takane *et al.* (2004) is indicated with dashed lines.

the three linked SNPs, −1847T>C, −1347T>C and −129T>C. This network profile supported the previous classification of promoter region haplotypes by Takane *et al.* (2004), as indicated by the dashed lines. However, our current study revealed the presence of additional subtypes. Detailed comparisons between our Block 1 haplotypes and previously described promoter region haplotypes are described in the next section.

The Block 2 cladogram showed that there were four major haplotypes, *2d, *10a, *1e and *8a, and that most of the minor haplotypes appeared to be derived from *1e, *2d or *8a (Fig. 6c). Network analysis showed that the *2 and *8 haplotypes, which share the SNP 1236C>T, were distant from *1e and *10a, and that *10a and the adjacent *13a that both have 2677G>A without the common polymorphisms 1236C>T, 2677G>T and 3435C>T, are relatively closely related to the *1 group. The *8 group bearing 1236C>T was highly diverse and included many haplotypes. The *4 (1236C>T and 3435C>T) and *9 (1236C>T and 2677G>T) groups were related to *8 and *2. This network profile supported the previous classification of Block 2 haplotypes based on common polymorphisms. For the *6 group, containing 3435C>T, *6a and *6b were distantly related to *1 haplotypes, which suggested that different mutational and/or recombinational events were responsible for these haplotypes.

For Block 3, *1b and the other minor *1 haplotypes were related to the major *1a haplotype, while the *2 group (*2a and *2b) with nonsynonymous SNPs appeared to be derived from either *1a or *1b (Fig. 6d).

Comparison of Block 1 Haplotypes with Reported Promoter Haplotypes

To compare our Block 1 haplotype structures with the previously reported promoter region haplotypes (Taniguchi *et al.* 2003: Takane *et al.* 2004), we classified Block 1 haplotypes into 6 subgroups (A, B, E, G, J and L) based on the marker sites in the reported haplotypes and the network analysis performed in our present study. The summary of this comparison is shown in Fig. 7.

Haplotypes that did not harbour any of the previous markers were classified into subgroup A, in which *1a was the major haplotype. The B subgroup, which included *1b, was defined as haplotypes that contained −1789G>A. Haplotypes that contained additional SNPs −371A>G or −145C>G were classified into subgroups J (including *1j) or L (including *1L), respectively. Subgroup G was defined as those haplotypes that contained the three linked variations (−1847T>C, −1347T>C and −129T>C) in which *1g was the major haplotype, and subgroup E (including *1e) was defined as those haplotypes with the three linked variations plus -1461_-1457delCATCC.

As indicated by the cladograms in the previous section, our study revealed that subgroup A, previously classified as wild-type, could be further classified into six types: the major *1a type without any marker variation and five other types with either IVS1 −78delG (*1c), IVS4 − 25G>T (*1d), 325G>A(E109K) (*2a), IVS5 + 123A>G (*1n), or 304 G>C(G102R) (*3a). Each of the B, J, and L subgroups that shared −1789G>A were further divided into two types based on the presence of IVS5 + 123A>G. Subgroup G, with the three reported marker SNPs, was also linked to IVS3 + 36C>T, and this subgroup was further characterized by the presence of IVS4 − 25G>T (*1h and *4a?) or IVS5 + 76T>G (*1g). Subgroup E was linked with both IVS3 + 36C>T and IVS5 + 76T>G.

Ethnic Differences

It is well known that there are differences in the frequencies of functionally important haplotypes involving common SNPs (1236C>T, 2677G>T/A and 3435C>T) and promoter region SNPs between different ethnic groups (Kim *et al.* 2001; Kroetz *et al.* 2003; Tang *et al.* 2002, 2004; Takane *et al.* 2004). To characterize these haplotypes in the Japanese population, we compared the frequencies of Block 1 and Block 2 haplotypes that harbour common SNPs with representative reported data from different ethnic groups (Kroetz *et al.* 2003; Takane *et al.* 2004) (Tables 4 and 5). Block 1 haplotype frequencies were generally consistent with those from previous reports for Japanese (Takane *et al.* 2004), except that our study did not detect the reported H7 haplotype that contained −1154T>C alone (Table 4). It has also been suggested that there is much more haplotypic variation in Japanese than in Caucasian populations (Takane *et al.* 2004), and our study supported this.

As for Block 2 haplotypes, the *1 and *2 groups were the common major haplotypes in all the ethnic groups.

Site	5'-Flanking					Ex. 1(5'-UTR)		Int. 1	Ex. 2	Int. 3	Ex. 4	Int. 4	Ex. 5		Int. 5	
Position[a]	-1847	-1789	-1451 /-1457	-1347	-371	-145	-129	IVS1 -78	49	IVS3 +36	144	IVS4 -25	334	385	IVS5 +76	IVS5 +123
Nucleotide change	T>C	G>A	del/CAT CC	T>C	A>G	C>G	T>C	del/G	T>C	C>T	G>T	C>T	G>C	G>A	T>G	A>G
Amino acid change									F17L		K86N		G107R	E129K		

Haplotypes		Tagging variations in the previous reports	Additional tagging variations in this study	No. of chromosome	Reported haplotype[j]	
Subgroup	Type				Taniguchi et al (2003)	Takane et al (2004)
A	*1a type[b]			559	H1	H1
	*1c type[c]		delG	88		
	*1d type[d]		T	48		
	*2a		delG 2	18		
	*1u		G	11		
	*3a		delG 3	5		
	minors[e]		(other combinations of SNPs)	5		
B	*1b type[f]	A	G	126	H2 (low)	H4
	*1f	A		30		
	*5a?	A	5 G	1		
J	*1j	A G		17	H5(nd)	H5
	*1k	A G	G	16		
L	*1l	A G		13	H2 (low) or H5 (nd)	H6 (low)
	*1m	A G	G	11		
	minors[g]	A (G) G	(other combinations of SNPs)	3		
G	*1q	C C G C	T G	30	H3 (low)	H2 (high)
	*1h	C C G C	T T	23		
	*4a?	C C G C	4 T T	1		
	minors[h]	C C G C	(other combinations of SNPs)	4		
E	*1e type[i]	C del C G C	T G	32		H3 (high)

Figure 7 New classification of Block 1 haplotypes and comparison with reported promoter region haplotypes. Genetic variations (allele frequency >0.01) and nonsynonymous variations in Block 1 were sorted according to marker variation, and classified into 6 subgroups (A, B, J, L, G and E).

[a]The positions in other reports were adjusted to the nucleotide numbers used in this study.

[b]The *1a type includes *1a, *1o, *1s, *1t, *1v, *1w, *1x, *1y.

[c]The *1c type includes the *1c haplotype and an ambiguously defined *1 haplotype.

[d]The *1d type includes the *1d haplotype and two ambiguously defined *1 haplotypes.

[e]"Minors" include the *1u and *1r haplotypes and one ambiguously defined *1 haplotype.

[f]The *1b type includes the *1b and *1i haplotypes and three ambiguously defined *1 haplotypes.

[g]"Minors" include the *1q haplotype and one ambiguously defined *1 haplotype.

[h]"Minors" include the *1p haplotype and one ambiguously defined *1 haplotype.

[i]The *1e type includes the *1e haplotype and one ambiguously defined *1 haplotype.

[j]Altered promoter activity in the reporter gene assay is shown in parenthesis.

nd; not determined.

However, the frequency of the *2 group was much lower than that of the *1 group in Africans. The frequencies of *4 and *8 were higher in Japanese than in Caucasians, and the frequency of the *6 group was higher in Caucasians than in other ethnic groups. The most prominent characteristic of the Japanese population was the high frequency of *10 compared with the other ethnic groups. The variations that characterized *11 to *18 were only detected in our study, probably due to the relatively large number of subjects used. The haplotype distribution in Japanese was similar to that described for Asians, but with slight differences in the frequencies of *6, *8, *9, and *10 reported for a mixed Asian population (Kroetz *et al.* 2003).

Tagging SNPs for *ABCB1* Genotyping

For genotyping *ABCB1* in association studies it would be critical to select SNPs for the major haplotypes, including functional ones in Blocks 1 and 2. Table 6 shows the major tagging SNPs for genotyping which are applicable to Japanese and also to other ethnic populations. Genotyping with these SNPs can assign the diplotypes of Blocks 1 and 2 in more than 95% of Japanese. The nonsynonymous SNPs in Blocks 1 and 2, and the additional tagging variations in Block 1 obtained in our study (Fig. 7), could be included in the list for evaluation of their functional significance.

Table 4 Ethnic differences in *ABCB1* Block 1 haplotypes

Marker site[a]	This study: Subgroup (see Fig. 7)	This study: Japanese (n = 1066)	Reported data (Takane *et al.* 2004): Group	Reported data (Takane *et al.* 2004): Japanese (n = 188)	Reported data (Takane *et al.* 2004): Caucasian (n = 192)
	A	0.712	H1	0.665	0.964
−1789G>A	B	0.147	H4	0.191	nd
−1789G>A, −371A>G	J	0.031	H5	0.027	nd
−1789G>A, −145C>G	L	0.025	H6	0.032	nd
−1847T>C[b]	G	0.054	H2	0.043	nd
−1461delCATCC, −371A>G, −1847T>C[b]	E	0.030	H3	0.037	nd
−1154T>C		nd	H7	0.005	nd
−1753delGA		nd	H8	nd	0.010
−1347T>C, −129T>C		nd	H9	nd	0.016
−1085A>G		nd	H10	nd	0.010

[a]Each reported position was adjusted to the nucleotide numbers used in this study.
[b]This SNP is linked to −1347T>C and −129T>C.
n; 2 × number of subjects.
nd; not detected.

Table 5 Ethnic differences in the *ABCB1* Block 2 haplotypes

Group	This study: Japanese (n = 1066)	Reported data (Kroetz *et al.* 2003)[a]: Asian (n = 60)	Caucasian (n = 200)	African (n = 200)
*1	0.216	0.216	0.370	0.721
*2	0.386	0.365	0.410	0.075
*3	nd	nd	0.010	0.010
*4	0.016	0.016	0.005	0.090
*6	0.034	0.016	0.120	0.035
*7	nd	nd	0.015	0.005
*8	0.141	0.216	0.010	0.040
*9	0.020	0.082	0.025	0.010
*10	0.174	0.066	0.025	0.005
*11	0.005	nd	nd	nd
*12	0.002	nd	nd	nd
*13	0.002	nd	nd	nd
*14	0.001	nd	nd	nd
*15	0.001	nd	nd	nd
*16	0.001	nd	nd	nd
*17	0.001	nd	nd	nd
*18	0.001	nd	0.01	nd

[a]Reported haplotypes were re-assigned according to our haplotype nomenclature.
n = 2 × number of subjects.
nd; not detected.

Discussion

Extensive studies of *ABCB1* haplotypes and their functional significance have been conducted, mostly focused on the common SNPs of 1236C>T, 2677G>T/A, and 3435C>T. However, recent association studies on promoter region haplotypes have indicated the importance of haplotypes within this region (Taniguchi *et al.* 2003; Takane *et al.* 2004). The results of functional or P-gp expression analyses based on these polymorphisms/haplotypes have not always been consistent, possibly due to the small number of subjects used, different ethnic backgrounds, or insufficient haplotyping over a limited region. In the present study, we have conducted a re-assignment of Block 1 haplotypes by extending the region sequenced to the distal promoter, and

Table 6 Major tagging SNPs of *ABCB1* for genotype-phenotyhpe association studeis

i) Block 1 haplotypes (subgroups)					
Position	−1847	−1789	−1461_−1457	−371	−145
Nucleotide change	T > C	G > A	delCATCC	A > G	C > G
	a	a	a	a	a
A					
B		A			
J		A		G	
L		A			G
G	C			G	
E	C		del	G	

ii) Block 2 haplotypes				
Position	1236		2677	3435
Nucleotide change	C > T	G > A	G > T	C > T
Amino acid change		A893T	A893S	
		a		
*1				
*2	T		T	T
*4	T			T
*6				T
*8	T			
*9	T		T	
*10		A		
*11	T	A		
*18		A		T

[a]Specific for Asian populatioins.

added novel haplotypes in other blocks after assessing a large number of subjects.

LD analysis revealed that one of the marker SNPs in the promoter region, −1789G>A, was moderately linked to IVS5 + 123A>G, previously classified into Block 2. Therefore, we shifted the border between Block 1 and Block 2 and re-analyzed the Block 1 haplotypes. Two promoter haplotype classes associated with functional changes have been reported previously (Taniguchi *et al.* 2003; Takane *et al.* 2004). One class included the −1789G>A SNP, and the other included the three linked SNPs of −1847T>C, −1347T>C and −129T>C. In our analysis these SNPs were included in our Block 1 region.

The haplotype containing −1789G>A was reported to be associated with reduced P-gp expression levels in the colon and liver, and reduced promoter activity was shown in a reporter gene assay (Taniguchi *et al.* 2003) (see Fig. 7). However, another study found that a haplotype containing −1789G>A without −145C>G (subgroups B and J in our present study) showed no change in the reporter assay, while another haplotype that contained −1789G>A together with −145C>G (subgroup L) showed reduced promoter activity (Takane *et al.* 2004). Data on the functional effects of haplotypes harbouring the three linked SNPs (G and E subgroups) are also contradictory. While one study showed an association with reduced colon and liver P-gp expression levels in patients and reduced promoter activity in a reporter gene assay (Taniguchi *et al.* 2003), another study reported an association with increased P-gp expression levels in the placenta and liver, and with increased promoter activity in a reporter gene assay (Takane *et al.* 2004). By expanding Block 1 into intron 5 we identified additional types within previously reported wild-type sequences (corresponding to subgroup A in this study) and other variant haplotypes (subgroups B, E, G, J, and L) (Fig. 7). In total our data revealed 11 tagging variations in Block 1: −1789G>A, −1461_−1457CATCdel, −371A>G, −145C>G, −129T>C, IVS1 − 78delG, IVS4 − 25G>T, 304G>C (G102R), 325G>A (E109K), IVS + 76T>G and IVS5 + 123A>G. Thus, if some of these markers are of functional importance it is possible that

our subdivisions (types within A and other subgroups) might explain the discrepancies in P-gp expression levels in the previously reported studies. In fact, our preliminary observation has suggested possible influences of some of the tagging variations in Block 1 on pharmacokinetic parameters of paclitaxel (data not shown). However, this hypothesis requires further clarification in large scale clinical studies.

Several novel haplotypes were added to the other 3 blocks (1, 38, and 4 new haplotypes in Blocks −1, 2, and 3, respectively). We identified a new haplotype *1d in Block −1, but this variant haplotype was very rare and the functional significance of uncommon Block −1 haplotypes remains unknown. We added 7 new groups to Block 2 haplotypes (*12 to *18), but their frequencies were also very low (0.002 and less). We also confirmed the previous finding that, in order of frequency, the major groups were *2d, *10a, *1e, and *8a. In our previous study we estimated the relative P-gp activity of the different haplotypes according to the renal clearance of irinotecan and its metabolites in Japanese cancer patients. While we found a significant association between *2, which contained the three common SNPs, and reduced renal clearance levels, associations with the *6, *8, and *10 groups that contained only one of the common markers remained unclear. For the *4, *9, and *11 groups, which harbour two marker SNPs in Block 2, functional evaluation was impossible due to the small number of subjects. Previously we showed that *1f may have been associated with reduced P-gp activity. The current study revealed that *1f in Block 2 was completely linked with the newly defined *1d in Block 1, which contained IVS4 −25G>T. A further association study is needed to clarify the effects of the linked *1d (Block 1) and *1f (Block 2) haplotypes. Regarding Block 3 we added several minor *1 haplotypes and confirmed the previous findings that *1a and *1b were the major haplotypes. We previously observed a trend for an association between *1b and higher P-gp activity. Taking into consideration the haplotype-combinations across the blocks this trend also needs to be confirmed in a larger number of subjects.

It is well recognized that there are large ethnic differences in the frequencies of functionally important haplotypes, including 1236C>T, 2677G>T, and 3435>T (corresponding to the *2 group in Block 2), and promoter region SNPs (corresponding to the variant Block 1 subgroups). Comparison of our data with the results from other ethnic groups indicated the existence of unique haplotype profiles in the Japanese population. As suggested by the previous report on the promoter region (Takane *et al.* 2004), Japanese samples exhibited large variations in Block 1 haplotypes. This suggested that not only *2 in Block 2 but also certain Block 1 haplotypes may be functionally important in the Japanese ethnic group. For Block 2 we confirmed our previous findings that the major groups were *1 and *2, and that *2d was the most frequent haplotype. While both groups were detected as the major types in other Asian and Caucasian populations, *1 was considerably more frequent than *2 in Africans (Kroetz *et al.* 2003). Another recent study found that the two major haplotypes were common to 5 ethnic groups (Tang *et al.* 2004). That study also revealed that the Chinese and Malay haplotype profiles were very similar, and that while some similarities were also observed between Caucasian and Indian populations, Africans differed from all other non-African populations. Furthermore, their study suggested that positive selection for 2677T-3435T had occurred in Chinese and Malays, and for 3435C in Africans. As pointed out previously, frequent occurrence of *10 (2677G>A) was unique to Japanese compared with Caucasians and African populations. Our study revealed higher frequencies of *10 (2677G>A) and *6 (3435C>T) and lower frequencies of *8 (1236C>T) and *9 (1236C>T and 2677G>T) than reported for Asian populations in Kroetz *et al.* (2003) (Table 5). This difference might be due to the mixed Asian population used in the report, as differences in the frequencies of 2677G>A between the Chinese, Malay, and Indian populations have been noted (Tang *et al.* 2004). The finding that the high frequency of 2677G>A is shared among Japanese, Koreans (Yi *et al.* 2004) and Chinese (Tang *et al.* 2004) suggests a close evolutionary relationship between these three populations.

A whole-genome haplotype database for three populations is now available at the Perlegen website (www.perlegen.com), which provides a good tool for investigation of the structures of human genetic variation within and between different populations (Hinds *et al.* 2005). For the *ABCB1* gene, however, we could not directly compare their data with ours because their

SNPs are mostly intronic and did not overlap with our SNP markers (<20%).

For genotype-phenotype association studies on the *ABCB1* gene, genotyping of the major functional key SNPs in Blocks 1 and 2 (Table 6) would be useful. Further studies on the clinical significance of the haplotypes described in the present study and elucidation of the haplotype-combinations across blocks, will be required to achieve the goal of personalized drug therapy.

Conclusions

We re-established *ABCB1* haplotypes in the Japanese population based on novel polymorphisms found in a large number of subjects, expanding the promoter region. Our current data added more detailed information on functionally-important haplotypes in Blocks 1 and 2 in the Japanese population, and identified differences in haplotype profiles between ethnic groups. The information provided in this study will be of use in further studies investigating the relationship between genetic markers and functional changes.

Acknowledgements

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

References

Chowbay, B., Cumaraswamy, S., Cheung, Y. B., Zhou, Q. & Lee, E. J. (2003) Genetic polymorphisms in MDR1 and CYP3A4 genes in Asians and the influence of MDR1 haplotypes on cyclosporin disposition in heart transplant recipients. *Pharmacogenetics* **13**, 89–95.

Cordon-Cardo, C., O'Brien, J. P., Casals, D., Rittman-Grauer, L., Biedler, J. L., Melamed, M. R. & Bertino, J. R. (1989) Multidrug-resistance gene (P-glycoprotein) is expressed by endothelial cells at blood-brain barrier sites. *Proc Natl Acad Sci U S A* **86**, 695–698.

Debry, P., Nash, E. A., Neklason, D. W. & Metherall, J. E. (1997) Role of multidrug resistance P-glycoproteins in cholesterol esterification. *J Biol Chem* **272**, 1026–1031.

Drach, J., Gsur, A., Hamilton, G., Zhao, S., Angerler, J., Fiegl, M., Zojer, N., Raderer, M., Haberl, I., Andreeff, M. & Huber, H. (1996) Involvement of P-glycoprotein in the transmembrane transport of interleukin-2 (IL-2), IL-4, and interferon-gamma in normal human T lymphocytes. *Blood* **88**, 1747–1754.

Fojo, A. T., Ueda, K., Slamon, D. J., Poplack, D. G., Gottesman, M. M. & Pastan, I. (1987) Expression of a multidrug-resistance gene in human tumors and tissues. *Proc Natl Acad Sci USA* **84**, 265–269.

Furuno, T., Landi, M. T., Ceroni, M., Caporaso, N., Bernucci, I., Napp, G., Martignoni, E., Schaeffeler, E., Eichelbaum, M., Schwab, M. & Zanger, U. M. (2002) Expression polymorphism of the blood-brain barrier component P-glycoprotein (MDR1) in relation to Parkinson's disease. *Pharmacogenetics* **12**, 529–534.

Geick, A., Eichelbaum, M. & Burk, O. (2001) Nuclear receptor response elements mediate induction of intestinal MDR1 by rifampin. *J Biol Chem* **276**, 14581–14587.

Goto, M., Masuda, S., Saito, H., Uemoto, S., Kiuchi, T., Tanaka, K. & Inui, K. (2002) C3435T polymorphism in the MDR1 gene affects the enterocyte expression level of CYP3A4 rather than Pgp in recipients of living-donor liver transplantation. *Pharmacogenetics* **12**, 451–457.

Hinds, D. A., Stuve, L. L., Nilsen, G. B., Halperin, E., Eskin, E., Ballinger, D. G., Frazer, K. A. & Cox, D. R. (2005) Whole-genome patterns of common DNA variation in three human populations. *Science* **307**, 1072–1079.

Hoffmann, U. & Kroemer, H. K. (2004) The ABC transporters MDR1 and MRP2: multiple functions in disposition of xenobiotics and drug resistance. *Drug Metab Rev* **36**, 669–701.

Hoffmeyer, S., Burk, O., von Richter, O., Arnold, H. P., Brockmoller, J., Johne, A., Cascorbi, I., Gerloff, T., Roots, I., Eichelbaum, M. & Brinkmann, U. (2000) Functional polymorphisms of the human multidrug-resistance gene: multiple sequence variations and correlation of one allele with P-glycoprotein expression and activity in vivo. *Proc Natl Acad Sci USA* **97**, 3473–3478.

Ieiri, I., Takane, H. & Otsubo, K. (2004) The MDR1 (ABCB1) gene polymorphism and its clinical implications. *Clin Pharmacokinet* **43**, 553–576.

Itoda, M., Saito, Y., Komamura, K., Ueno, K., Kamakura, S., Ozawa, S. & Sawada, J. (2002) Twelve Novel Single Nucleotide Polymorphisms in ABCB1/MDR1 among Japanese Patients with Ventricular Tachycardia who were Administered Amiodarone. *Drug Metab Pharmacokinet* **17**, 566–571.

Johne, A., Kopk, K., Gerloff, T., Mai, I., Rietbrock, S., Meisel, C., Hoffmeyer, S., Kerb, R., Fromm, M. F., Brinkmann, U., Eichelbaum, M., Brockmoller, J., Cascorbi, I. & Roots, I. (2002) Modulation of steady-state kinetics of digoxin by haplotypes of the P-glycoprotein MDR1 gene. *Clin Pharmacol Ther* **72**, 584–594.

Kim, R. B., Leake, B. F., Choo, E. F., Dresser, G. K., Kubba, S. V., Schwarz, U. I., Taylor, A., Xie, H. G., McKinsey, J., Zhou, S., Lan, L. B., Schuetz, J. D., Schuetz, E. G. & Wilkinson, G. R.. (2001) Identification of functionally

variant MDR1 alleles among European Americans and African Americans. *Clin Pharmacol Ther* **70**, 189–199.

Kim, R. B. (2002) MDR1 single nucleotide polymorphisms: multiplicity of haplotypes and functional consequences. *Pharmacogenetics* **12**, 425–427.

Kimchi-Sarfaty, C., Gribar, J. J. & Gottesman, M. M. (2002) Functional characterization of coding polymorphisms in the human MDR1 gene using a vaccinia virus expression system. *Mol Pharmacol* **62**, 1–6.

Kitamura, Y., Moriguchi, M., Kaneko, H., Morisaki, H., Morisaki, T., Toyama, K. & Kamatani, N. (2002) 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.

Kurata, Y., Ieiri, I., Kimura, M., Morita, T., Irie, S., Urae, A., Ohdo, S., Ohtani, H., Sawada, Y., Higuchi, S. & Otsubo, K. (2002) Role of human MDR1 gene polymorphism in bioavailability and interaction of digoxin, a substrate of P-glycoprotein. *Clin Pharmacol Ther* **72**, 209–219.

Kroetz, D. L., Pauli-Magnus, C., Hodges, L. M., Huang, C. C., Kawamoto, M., Johns, S. J., Stryke, D., Ferrin, T.E., DeYoung, J., Taylor, T., Carlson, E. J., Herskowitz, I., Giacomin, i K. M. & Clark, A. G. (2003) Pharmacogenetics of Membrane Transporters Investigators. Sequence diversity and haplotype structure in the human ABCB1 (MDR1, multidrug resistance transporter) gene. *Pharmacogenetics* **13**, 481–494.

Liu, Y. Y., Han, T. Y., Giuliano, A. E. & Cabot, M. C. (2001) Ceramide glycosylation potentiates cellular multidrug resistance. *FASEB J* **15**, 719–730.

Morita, N., Yasumori, T. & Nakayama, K. (2003) Human MDR1 polymorphism: G2677T/A and C3435T have no effect on MDR1 transport activities. *Biochem Pharmacol* **65**, 1843–1852

Riordan, J. R., Deuchars, K., Kartner, N., Alon, N., Trent, J. & Ling, V. (1985) Amplification of P-glycoprotein genes in multidrug-resistant mammalian cell lines. *Nature* **316**, 817–819.

Sai, K., Kaniwa, N., Itoda, M., Saito, Y., Hasegawa, R., Komamura, K., Ueno, K., Kamakura, S., Kitakaze, M., Shirao, K., Minami, H., Ohtsu, A., Yoshida, T. Saijo, N., Kitamura, Y., Kamatani, N., Ozawa, S. & Sawada, J. (2003) Haplotype analysis of ABCB1/MDR1 blocks in a Japanese population reveals genotype-dependent renal clearance of irinotecan. *Pharmacogenetics* **13**, 741–757.

Schwab, M., Schaeffeler, E., Marx, C., Fromm, M. F., Kaskas, B., Metzler, J., Stange, E., Herfarth, H., Schoelmerich, J., Gregor, M., Walker, S., Cascorbi, I., Roots, I., Brinkmann, U., Zanger, U. M. & Eichelbaum, M. (2003) Association between the C3435T MDR1 gene polymorphism and susceptibility for ulcerative colitis. *Gastroenterology* **124**, 26–33.

Siegsmund, M., Brinkmann, U., Schaffeler, E., Weirich, G., Schwab, M., Eichelbaum, M., Fritz, P., Burk, O., Decker, J., Alken, P., Rothenpieler, U., Kerb, R. & Hoffmeyer, S. (2002) Association of the P-glycoprotein transporter MDR1(C3435T) polymorphism with the susceptibility to renal epithelial tumors. *J Am Soc Nephrol* **13**, 1847–1854.

Takane, H., Kobayashi, D., Hirota, T., Kigawa, J., Terakawa, N., Otsubo, K. & Ieiri, I. (2004) Haplotype-oriented genetic analysis and functional assessment of promoter variants in the MDR1 (ABCB1) gene. *J Pharmacol Exp Ther* **311**, 1179–1187.

Tanabe, M., Ieiri, I., Nagata, N., Inoue, K., Ito, S., Kanamori, Y., Takahashi, M., Kurata, Y., Kigawa, J., Higuchi, S., Terakawa, N. & Otsubo, K. (2001) Expression of P-glycoprotein in human placenta: relation to genetic polymorphism of the multidrug resistance (MDR)-1 gene. *J Pharmacol Exp Ther* **297**, 1137–1143.

Tang, K., Ngoi, S. M., Gwee, P. C., Chua, J. M., Lee, E. J., Chong, S. S. & Lee, C. G. (2002) Distinct haplotype profiles and strong linkage disequilibrium at the MDR1 multidrug transporter gene locus in three ethnic Asian populations. *Pharmacogenetics* **12**, 437–450.

Tang, K., Wong, L. P., Lee, E. J., Chong, S. S. & Lee, C. G. (2004) Genomic evidence for recent positive selection at the human MDR1 gene locus. *Hum Mol Genet* **13**, 783–797.

Taniguchi, S., Mochida, Y., Uchiumi, T., Tahira, T., Hayashi, K., Takagi, K., Shimada, M., Maehara, Y., Kuwano, H., Kono, S., Nakano, H., Kuwano, M. & Wada, M. (2003) Genetic polymorphism at the 5′ regulatory region of multidrug resistance 1 (MDR1) and its association with interindividual variation of expression level in the colon. *Mol Cancer Ther* **2**, 1351–1359.

van Helvoort, A., Smith, A. J., Sprong, H., Fritzsche, I., Schinkel, A. H., Borst, P. & van Meer, G. (1996) MDR1 P-glycoprotein is a lipid translocase of broad specificity, while MDR3 P-glycoprotein specifically translocates phosphatidylcholine. *Cell* **87**, 507–517.

Wong, M., Evans, S., Rivory, L. P., Hoskins, J. M., Mann, G. J., Farlow, D., Clarke, C. L., Balleine, R. L. & Gurney, H. (2005) Hepatic technetium Tc 99m-labeled sestamibi elimination rate and ABCB1 (MDR1) genotype as indicators of ABCB1 (P-glycoprotein) activity in patients with cancer. *Clin Pharmacol Ther* **77**, 33–42.

Yi, S. Y., Hong, K. S., Lim, H. S., Chung, J. Y., Oh, D. S., Kim, J. R., Jung, H. R., Cho, J. Y., Yu, K. S., Jang, I. J. & Shin, S. G. (2004) A variant 2677A allele of the MDR1 gene affects fexofenadine disposition. *Clin Pharmacol Ther* **76**, 418–427.

Received: 20 April 2005
Accepted: 21 October 2005