RACIAL VARIABILITY IN HAPLOTYPE FREQUENCIES OF UGT1A1 AND GLUCURONIDATION ACTIVITY OF A NOVEL SINGLE NUCLEOTIDE POLYMORPHISM 686C>T (P229L) FOUND IN AN AFRICAN-AMERICAN

Nahoko Kaniwa, Kouichi Kurose, Hideto Jinno, Toshiko Tanaka-Kagawa, Yoshiro Saito, Mayumi Saeki, Jun-ichi Sawada, Masahiro Tohkin, and Ryuichi Hasegawa

Division of Medicinal Safety Science (N.K., K.K., M.T., R.H.), Division of Environmental Chemistry (H.J., T.T-K.), and Project Team for Pharmacogenetics (Y.S., M.S., J.S.), National Institute of Health Sciences, Tokyo, Japan

Received August 10, 2004; accepted November 23, 2004

ABSTRACT:
Ethnic differences in genetic polymorphisms in UDP-glucuronosyltransferase 1A1 (UGT1A1) were investigated among African-Americans, Caucasians, and Japanese using samples obtained from 150 individuals for each population. Genotyping of −3279T>G in the phenobarbital-responsive enhancer module, TA repeats in the TATA box, 211G>A (G71R) and 686C>A (P229Q) in exon 1, and three single nucleotide polymorphisms (SNPs) (1813C>G, 1941C>G, and 2042C>G) in the 3′-untranslated region in exon 5 was performed. Eight haplotypes of block 1 (exon 1 and its 5′-flanking region) harboring the first four variations were assigned to each individual. The dominant haplotype for African-Americans was *28b (−3279G;TAp; 211G;686C) (0.446), whereas that for the Japanese was *1a (−3279T; TAap;211G;686C) (0.610). Frequencies of the two haplotypes *1a and *28b were comparable in Caucasians. Haplotype *6a (−3279T;TAap; 211A;686C) was characteristic of the Japanese, whereas haplotypes *36b and *37b (−3279T;TAap and TAap;211G;686C) were found mostly in African-Americans. Although the three SNPs in block 2 (exons 2–5) were in complete linkage in the Japanese, they were not completely linked in African-Americans or Caucasians. These differences in haplotype distribution patterns among the three populations suggest the possibility of ethnic differences in toxicity profiles of drugs detoxicated by UGT1A1. A novel SNP, 686C>T (P229L), was found in an African-American. The intrinsic clearance of 7-ethyl-10-hydroxycamptothecin (SN-38) by P229L UGT1A1 expressed in COS-1 cells was about 3% of the wild type. The results of Western blotting and real-time reverse transcription-polymerase chain reaction suggest that the low glucuronidation activity of the variant was partly due to its low stability. The variation 686C>T may cause high toxicity during 7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxycamptothecin (CPT-11) therapy or hyperbilirubinemia in patients.

The wide interindividual variability in SN-38G formation in hepatic tissues is known and has been shown to correlate with UGT1A1 genetic polymorphisms (Iyer et al., 1999). To date, a wide variety of allelic polymorphisms of human UGT1A1 have been reported, most of which lead to a partial or complete deficiency of enzyme activity and are associated with diseases, such as the Crigler-Najjar (CN) syndrome types I and II or Gilbert’s syndrome, with hyperbilirubinemia and jaundice as main symptoms (http://som.flinders.edu.au/FUSA/ClinPharm/UGT1A1alleles.html; Mackenzie et al., 1997; Tukey and Strassburg, 2000). To understand these syndromes and their possible genetic causes, polymorphisms in the TATA box have been the most extensively studied. An inverse relationship between the number of TA repeats (five to eight) and transcriptional activity has been reported. RT-PCR, reverse transcription-polymerase chain reaction.

ABBREVIATIONS: SN-38, 7-ethyl-10-hydroxycamptothecin; CPT-11, 7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxycamptothecin; SN-38G, 7-ethyl-10-hydroxycamptothecin glucuronide; CN, Crigler-Najjar; SNP, single nucleotide polymorphism; PCR, polymerase chain reaction; RT-PCR, reverse transcription-polymerase chain reaction.
RACIAL VARIABILITY IN HAPLOTYPE FREQUENCIES OF UGT1A1

2000; Iyer et al., 2002). Other variations often associated with neonatal hyperbilirubinemia, CN syndromes, Gilbert’s disease, or irinotecan toxicity are 211G>A (G71R), 1456T>G (Y486D), and 686C>A (P229Q) (Ando et al., 2000; Kraemer and Klinker, 2002; Yamamoto et al., 2002) (In this report, the A of the translational start codon of the cDNA is designated 1.) Recently, the possibility of a synergistic effect of 3′-untranslated region variations with −3279T>G on the bilirubin level and SN-38 pharmacokinetics was also proposed (Sai et al., 2004). The single nucleotide polymorphism (SNP) −3279T>G is located in the phenobarbital-responsive enhancer module, which is involved in activation of UGT1AI transcription by the constitutive androstane receptor (Sugatani et al., 2002).

A recent pharmacogenetic study has suggested that there are advantages in the use of haplotypes rather than individual SNPs to investigate the association between genotypes and phenotypes (Judson et al., 2000). Racial variability in haplotype frequencies of UGT1A1 may contribute to an interpretation of the racial diversity of pharmacokinetics and toxicity of drugs metabolized by the enzyme. Recently, Sai et al. (2004) revealed the haplotype structure of UGT1A1 in Japanese cancer patients using −3279T>G, TA repeats in the TATA box, 211G>A, 686C>A, and three SNPs (1813C>T, 1941C>G, and 2042C>G) in the 3′-untranslated region in exon 5 as markers. Therefore, we studied in this report the differences in the haplotype frequencies of UGT1A1 among African-American, Caucasian, and Japanese populations using genomic DNA obtained from 150 individuals in each population, assessing the genotypes of these markers with dideoxy sequencing and pyrosequencing methods previously reported by Saeki et al. (2003). The function of a novel SNP 686C>T (P229L) that was found in an African-American individual during this study was also investigated by utilizing a heterologous expression system with COS-1 cells.

Materials and Methods

Materials

Materials, SN-38 (lot 970507R) and SN-38G (lot 970507R) were kindly supplied by Yakult Honsha Co. Ltd. (Tokyo, Japan). COS-1 cells were obtained from the Health Science Research Resources Bank (Osaka, Japan). Other chemicals used were all reagent grade.

Peripheral Blood Samples

One hundred fifty peripheral blood samples were obtained from healthy Japanese volunteers, 100 of which were kindly provided by Professor Ichiro Ieiri from the Tottori University Faculty of Medicine. Written informed consent was obtained from all participants. Peripheral blood samples of healthy Caucasian and African-American volunteers (150 each) were purchased from the Tennessee Blood Service Corporation (Memphis, TN). The ethics committee of the National Institute of Health Sciences approved this study.

UGT1A1 Genotyping by Pyrosequencing and Dideoxy Sequencing

Genotyping of UGT1A1 using Pyrosequence methods was performed as previously reported (Saeki et al., 2003). Briefly, genomic DNA (10–15 ng) was amplified by Ex-Tag (1 U; Takara Shuzo, Osaka, Japan) with a specific primer pair in which one of the primers was biotinylated. The primers for amplification and sequencing used for detection of polymorphisms −3279T>G, the TA repeat, 211G>A (G71R), and 686C>A/T (P229Q/L) have been shown previously (Saeki et al., 2003). Genotyping of 1941C>G for the Japanese samples was performed using an amplification primer pair (biotin-ATTGTAATGTTATCCTGCCC and CATTACCTATTTTACTACT) and a sequencing primer (CAGTGGGCGCAG). Generation of the single-stranded fragment and annealing of the sequencing primers have been described previously (Saeki et al., 2003). Genotypes were determined using the PSQ 96MA (Biostage AB, Uppsala, Sweden) and the PSQ 96 SNP reagent set (Biostage AB). Genotyping of 1813C>T, 1941C>G, and 2042C>G for African-American and Caucasian samples as well as the detection of a genotype 686C>T in the African-American individual were done by dideoxy sequencing as reported previously (Saeki et al., 2002; Sai et al., 2004).

Haplotype Analysis

Genotyping was successfully determined for 150 Japanese, 147 Caucasians, and 148 African-Americans. Dipolettes (combinations of haplotypes) in both blocks 1 and 2 were inferred separately for each population by PHASE version 2.0 (Stephens et al., 2001; Stephens et al., 2003).

Expression and Enzyme Assay of Wild-Type and Variant UGT1A1s

Expression of wild-type and variant UGT1A1s in COS-1 cells was carried out as described previously (Jinno et al., 2003b) with a minor modification; attB-flanked UGT1A1 cDNA was amplified by the two-step attB adaptor polymerase chain reaction (PCR) (Jinno et al., 2003a) using pcDNA3.1-UGT1AI/WT (Jinno et al., 2003b) as a template and gene-specific primers, 5′-AAAAAGCAGGCTGCAAAGGCGCCATGGCTGT-3′ and 5′-AGAAAGCCTGGTCCTACATGGTGCTTGGG-3′. The resulting PCR fragment was cloned into the pDONR201 vector (Invitrogen, Carlsbad, CA). The 686C>T mutation was introduced into the wild-type UGT1A1 cDNA clone in pDONR201, using a QuickChange multisite-directed mutagenesis kit (Stratagene, La Jolla, CA) with the 5′-phosphorylated oligonucleotide primer 5′-phospho-GGCACGTTTTATTCCTGGTTATGCAACACCTTGCTC-3′ (the nucleotide base is changed underlined). Subcloning of each UGT1AI fragment from pDONR201 into pcDNA-DEST40, a mammalian expression vector, was performed by the Gateway LR reaction.

Expression levels of UGT1AI mRNA and protein were determined as described previously (Jinno et al., 2003b). SN-38 glucuronidation activity of the wild-type and P229L variant UGT1AI was assayed according to the method of Hanioka et al. (2001).

Data Analysis

The χ² test in the SAS Preclinical Package (SAS Institute Japan Ltd., Tokyo, Japan) was applied to s × r contingency tables for comparison of genotype frequencies and analysis of haplotype distribution patterns among the three ethnic groups. In vitro Michaelis-Menten kinetic parameters were estimated by nonlinear regression analysis using Prism version 3.0 (GraphPad Software Inc., San Diego, CA). The average values of the kinetic parameters were calculated using results from three independent preparations. The t test supplied in Prism version 3.0 was applied to the comparison of the average values of protein expression and mRNA levels between wild-type and variant UGT1AI at a significance level of 0.05.

Results

Differences in Allele Frequencies among Three Ethnic Groups

Allele frequencies of polymorphisms in UGT1AI tested in this study are summarized in Table 1. The estimated allele frequency of −3279T>G (UGT1AI*60; for UGT1AI allele nomenclature refer to Mackenzie et al., 1997 and http://som.flanders.edu/AUSA/ClinPharm/UGT1AIalleles.htm) was 0.847, 0.550, and 0.257 in African-Americans, Caucasians, and Japanese, respectively, and they were significantly different (p < 0.0001 by the χ² test). Regarding the variation in the TATA box, the allele TA₅ was most common in the Japanese population, whereas frequencies of TA₅ and TA₆ (UGT1AI*28) were equally frequent in African-Americans. The allele frequencies of TA₅ (UGT1AI*36) or TA₆ (UGT1AI*37) in African-Americans were about 5%, but no carriers with TA₅ or TA₆ were found in the Japanese. The allele distribution patterns for the wild-type and three variants in the TATA box of the three ethnic groups were significantly different (p < 0.0001 by the χ² test). The variant 211G>A (G71R) (UGT1AI*6) was frequently found in the Japanese population, but not in the Caucasians population (only two heterozygous subjects), and none were found in African-Americans. The variant 686C>A (P229Q) (UGT1AI*27) was very rare in all the ethnic groups, and only one heterozygous carrier was found in the Japanese group. A novel variation in the same position as the variant 686C>A, but with a different nucleotide change (686C>T), was found in an African-American. This novel variant led to an amino acid change (686C>T), was found in an

Downloaded from utils.aspetjournals.org at ASPET Journals on June 23, 2017
<table>
<thead>
<tr>
<th>Position</th>
<th>Change</th>
<th>African-American</th>
<th>Caucasian</th>
<th>Japanese</th>
<th>( p ) (( \chi^2 ) test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of Heterozygotes/ Homozygotes</td>
<td>Allele Frequency (lower limit, upper limit)</td>
<td>Number of Heterozygotes/ Homozygotes</td>
<td>Allele Frequency (lower limit, upper limit)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(lower limit, upper limit)</td>
<td></td>
<td>(lower limit, upper limit)</td>
<td></td>
</tr>
<tr>
<td>(-3279)</td>
<td>T &gt; G</td>
<td>36/109 (0.386, 0.888)</td>
<td>65/30 (0.494, 0.606)</td>
<td>55/11 (0.208, 0.306)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TA box</td>
<td>T(_A_a), T(_A_B), T(_A_c)</td>
<td>17/1 (149)</td>
<td>20 (150)</td>
<td>0.097 (0.064, 0.130)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>211 G &gt; A G71R</td>
<td>0/0 (150)</td>
<td>2 (150)</td>
<td>35/6 (0.116, 0.198)</td>
</tr>
<tr>
<td></td>
<td>C &gt; A</td>
<td>686 C &gt; A P229Q</td>
<td>0/0 (150)</td>
<td>0 (150)</td>
<td>0.003 (0.000, 0.009)</td>
</tr>
<tr>
<td></td>
<td>C &gt; T</td>
<td>1813 C &gt; T P229L</td>
<td>0/1 (149)</td>
<td>0 (150)</td>
<td>0 (150)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1941 C &gt; G</td>
<td>43/6 (150)</td>
<td>47 (150)</td>
<td>25/2 (0.064, 0.130)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2042 C &gt; G</td>
<td>68/23 (150)</td>
<td>60 (150)</td>
<td>0.240 (0.192, 0.288)</td>
</tr>
</tbody>
</table>

* Values in parentheses are numbers of individuals determined.

* Assumed to be equal to the frequency for GT 1941 C > G.
2042C>G in African-Americans and Caucasians. The frequencies of 1813C>T and 2042C>G in the Japanese were not determined in this study, because these alleles were reported to be in complete association with 1941C>G in a Japanese population without exception (Sai et al., 2004; unpublished data). Therefore, allele frequencies of 1813C>T and 2042C>G in the Japanese were assumed to be equal to that of 1941C>G. On the other hand, we found that 1813C>T and 2042C>G were not always in association with 1941C>G in African-Americans and Caucasians, as shown in Table 1. The allele frequencies of the three SNPs were highest in African-Americans and lowest in the Japanese. Significant differences in the allele frequencies were detected \( (p < 0.0001, p = 0.0074, \text{and } p < 0.0001 \text{ for } 1813C>T, 1941C>G, \text{and } 2042C>G, \text{respectively, by the } \chi^2 \text{ test}).

**Differences in Haplotype Frequencies.** Sai et al. (2004) previously reported that *UGT1A1* could be divided into two blocks, with the transcription-regulating and promoter regions and exon 1 in block 1, and exons 2 to 5 in block 2, according to linkage disequilibrium analysis. Therefore, haplotype/diplotype analysis was performed using four marker variations in block 1 \( (-3279T>G, 211G>A, \text{and } 686C>A, \text{and the TATA box}), \text{and three marker variations in block 2 (1813C>T, 1941C>G, \text{and } 2042C>G). The diplotype configurations (combinations of haplotypes) were estimated for each subject by PHASE software (Stephens et al., 2001; Stephens et al., 2003). Concerning block 1, the diplotype configurations were inferred with greater than 0.99 certainties for 145 Japanese, 144 Caucasians, and 147 African-Americans. The haplotypes identified are summarized in Table 2 along with their frequencies for each population, where the 14 subjects with ambiguous diplootypes were excluded. We followed the haplotype nomenclature previously reported (Sai et al., 2004). One Japanese subject carrying heterozygous 686C>A also carried homozygous TA7. Thus, the association of variation 686C>A with TA7 was also confirmed in this study, as previously reported by Huang et al. (2000) and Sai et al. (2004). A new haplotype, designated *6d, having variations in both positions –3279 and 211, was identified that was not found in the previous study (Sai et al., 2004). We could not determine whether the novel variation 686T and the *28 allele were on the same chromosome, because the subject with 686T carried heterozygous TA repeats, TA6/TAD.

In block 1, haplotype *1A was predominant in the Japanese population \( (0.61 \text{ in frequency)\). In contrast, its frequencies in Caucasians and African-Americans were 0.45 and 0.15, respectively. Major haplotypes in African-Americans were *28b and *60a. Haplotype frequencies of *1A and *28b in Caucasians were comparable. Thus, the haplotype distribution patterns in block 1 for the individual populations were significantly different \( (p < 0.0001 \text{ by the } \chi^2 \text{ test}).

In block 2, the diplotype configurations using the three SNPs were inferred with certainties greater than 0.98 for all Caucasians and African-Americans (150 each). The haplotypes and their frequencies are summarized in Table 3. Four novel haplotypes, *IC, *ID, *IE, and *IF, were identified in the Caucasian and African-American populations, although *IE was not found in the latter. These haplotypes were not found in the Japanese, as shown in a previous study (Sai et al., 2004). Although the three SNPS were reported to be completely associated with each other in the Japanese (Sai et al., 2004), they were not always linked with each other in the rest of the populations. However, it is noteworthy that 1941G allele was always associated with 1813T and 2042G alleles except for the haplotype *IE, found only in two Caucasians. Haplotype *IA was predominant for all ethnic groups. The second major haplotype was *IB in both the Japanese and Caucasians. However, the frequencies of *IB and *IC were similar in African-Americans (0.183 and 0.163, respectively). The haplotype *IC was also found in Caucasians (0.06). Thus, the haplotype distribution patterns in block 2 for the three populations were also significantly different \( (p < 0.0001 \text{ by the } \chi^2 \text{ test}).

**Expression and SN-38 Glucuronidation of a Novel Variant (P229L) Compared with Wild-Type UGT1A1.** The relative expres-
sion level of the novel variant UGT1A1/P229L in the membrane fraction of COS-1 cells was determined by Western blotting using a polyclonal anti-human UGT1A antibody (Fig. 2). Because two bands, one at the same position as the wild-type and the other at a higher molecular weight, were detected for P229L, the total chemiluminescence of the two bands was used to calculate the protein expression level for P229L. The protein expression level of P229L was approximately 60% of the wild type, and this difference was statistically significant by the \( t \) test (\( p = 0.044 \)). The \( \text{UGT1A1} \) mRNA expression level was then measured by real-time reverse transcription-PCR (RT-PCR) using SYBR Green. As shown in Fig. 3, however, no significant difference in the mRNA level was detected among the COS-1 cells transfected with the expression plasmids carrying wild-type and P229L \( \text{UGT1A1} \) cDNAs.

The glucuronidation activity of SN-38 by P229L expressed in COS-1 cells was compared with that of the wild-type under 11 substrate concentrations ranging from 2.5 to 150 M. The representative curves of the Michaelis-Menten kinetics are shown in Fig. 4, and the estimated apparent kinetic parameters (\( K_m \), \( V_{max} \) and \( V_{max}/K_m \)) are summarized in Table 4. \( V_{max} \) values normalized

### Table 3

Haplotypes in block 2 (exons 2–5) of \( \text{UGT1A1} \) for three ethnic groups

<table>
<thead>
<tr>
<th>Position</th>
<th>1813</th>
<th>1941</th>
<th>2042</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleotide</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>G</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1A</td>
<td>0.617</td>
<td>0.757</td>
<td>0.903</td>
<td></td>
</tr>
<tr>
<td>*1B</td>
<td>0.183</td>
<td>0.157</td>
<td>0.097</td>
<td></td>
</tr>
<tr>
<td>*1C</td>
<td>0.163</td>
<td>0.060</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>*1D</td>
<td>0.033</td>
<td>0.017</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>*1E</td>
<td>0.000</td>
<td>0.007</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>*1F</td>
<td>0.003</td>
<td>0.003</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

*Haplotypes in Japanese were determined according to the genotype of 1941C > G, because it was previously reported that all three markers were completely associated in a Japanese population.

### Table 2

Haplotypes in block 1 (the enhancer/promoter regions and exon 1) of \( \text{UGT1A1} \) for three ethnic groups

<table>
<thead>
<tr>
<th>Position</th>
<th>-3279</th>
<th>TA box</th>
<th>211</th>
<th>686</th>
<th>Haplotype frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleotide</td>
<td>T</td>
<td>G</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*1a</td>
<td>0.150</td>
<td>0.451</td>
<td>0.610</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*6a</td>
<td>0.000</td>
<td>0.000</td>
<td>0.141</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*6d</td>
<td>0.000</td>
<td>0.000</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*28b</td>
<td>0.446</td>
<td>0.389</td>
<td>0.097</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*28c</td>
<td>0.000</td>
<td>0.000</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*36b</td>
<td>0.044</td>
<td>0.017</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*37b</td>
<td>0.065</td>
<td>0.007</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*60a</td>
<td>0.296</td>
<td>0.135</td>
<td>0.145</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Haplotypes in Japanese were determined according to the genotype of 1941C > G, because it was previously reported that all three markers were completely associated in a Japanese population.
FIG. 2. Expression of wild-type and variant P229L UGT1A1 in COS-1 cells. A, an aliquot (20 μg) of the pooled membrane fractions from three independent preparations was subjected to SDS-polyacrylamide gel electrophoresis, electrophoretically transferred to a polyvinylidene difluoride membrane, and immunochemically detected with a rabbit anti-human UGT1A antiserum. The membrane was subsequently stripped and reprobed with a rabbit anti-calnexin antiserum to confirm that the samples were evenly loaded. B, each Western blot from three independent preparations was quantified densitometrically, and the expression level of UGT1A1 proteins was normalized to that of the wild type. The results are expressed as the mean ± S.E.M. from three independent preparations. The expression level of UGT1A1 proteins of P229L was significantly lower than that of the wild type (p = 0.044).

FIG. 3. Quantification of UGT1A1 mRNA by real-time SYBR Green RT-PCR in COS-1 cells transfected with wild-type and variant P229L UGT1A1. UGT1A1 mRNA from the cellular total RNA samples was quantified by SYBR Green RT-PCR. Each sample was normalized on the basis of the β-actin content and expressed as a percentage of the wild type. The results indicate the mean ± S.E.M. from three independent preparations. No significant difference in mRNA level was observed between the wild type and P229L (p = 0.180).

to the expression levels are also shown. Wild-type UGT1A1 catalyzed SN-38 glucuronidation with an average apparent K_m value of 8.67 μM, whereas P229L catalyzed SN-38 glucuronidation with a K_m value of 37.6 μM. The average V_max values were 71.2 and 5.27 pmol/min/mg membrane protein for the wild-type and P229L, respectively. The average intrinsic clearances of SN-38 by glucuronidation (V_max/K_m) normalized to expressed UGT1A1 protein levels were 8.26 and 0.24 μl/min/mg protein for the wild-type and P229L, respectively.

Discussion

There have been previous reports of the allele frequencies of variants in the TATA box in various ethnic groups (Beutler et al., 1998; Fertrin et al., 2002; Innocenti et al., 2002; Sugatani et al., 2002; Ki et al., 2003; Premawardhena et al., 2003; Sai et al., 2004). According to these studies, allele frequencies of the wild-type TA_6 were highest in Polynesian people (0.97–0.99), followed by Asians, including Japanese, Koreans, and Taiwanese (0.81–0.92), and Caucasians (0.61–0.73). Allele frequencies were lowest in Africans (0.44–0.52). The allele frequencies of TA_6 and TA_8 in Africans were reported as approximately 0.05, but were very rare in Caucasians and were not detected in Asians. The allele frequencies of the four genotypes in the TATA box in African-American, Caucasian, and Japanese populations observed in this study were in the ranges previously reported. Allele frequencies of other UGT1A1 variations have not been as extensively studied as those in the TATA box. The allele frequencies for the variant −3279T>G were 0.17 to 0.27 in Japanese and Koreans (Sugatani et al., 2002; Ki et al., 2003; Sai et al., 2004), 0.47 in Caucasians (Innocenti et al., 2002), and 0.85 in African-Americans (Innocenti et al., 2002). Thus, our data (0.257 for Japanese, 0.550 for Caucasians, and 0.847 for African-Americans) are comparable with previously reported frequencies of this allele. The allele frequency of the variant 211G>A observed in this study was comparable to previously reported values for Asian people (0.09–0.21) (Huang et al., 2000; Sugatani et al., 2002; Ki et al., 2003; Sai et al., 2004). The variant is considered to be characteristic for Asians. We found only two individuals with this allele out of 150 Caucasians, and no African-Americans had this allele.

Innocenti et al. (2002) reported on the haplotype frequencies of UGT1A1 in African-Americans and Caucasians using −3279T>G, the TA repeat, and other SNPs in the transcription-regulating region. Although they did not use the marker SNPs 211G>A and 686C>A (likely due to their rarity in both populations), frequencies of the *28, *36, *37, and *60 haplotype groups in African-Americans can be calculated according to their data as 0.35, 0.04, 0.12, and 0.33, and the calculated frequencies in Caucasians as 0.36, 0.01, 0.01, and 0.09. These frequencies were also comparable to our data. We found two individuals carrying 211G>A in the Caucasian group; however, we could not infer haplotypes for these individuals with high certainty. The observed haplotype *28 group frequency in our study of 0.100 (*28b plus *28c) was slightly lower than the reported values of 0.131 in Japanese patients (Sai et al., 2004) and 0.127 in Koreans (Ki et al., 2003). However, the previously reported values are included in the 95% confidence interval of our data. The observed haplotype *6 group
frequency (∗6a plus ∗6b) was also comparable to the previously reported data in Japanese and Koreans. A strong association of TA7 with −3279G was suggested for Japanese cancer patients (Sai et al., 2004). This association was also observed in all populations measured in this study. Other variations, TA5 and TA6 in the TATA box, also had linkage with −3279G.

Recent studies indicate that 3′-untranslated regions may regulate mRNA stability (Day and Tuite, 1998; Conne et al., 2000). Exon 1 is unique for each member of the UGT1A subfamily, whereas exons 2 to 5 are common to all members of the subfamily. The variants in the 3′-untranslated region of UGT1A1 in exon 5, therefore, could have various effects on all enzymes of the UGT1A subfamily. However, until recently, their physiological effects have remained unknown. Acuna et al. (2003) reported a protective effect of the variant 1941C>G on liver transaminase levels caused by tolcapone, for which glucuronidation (mainly by UGT1A9) is one metabolic pathway (Acuna et al., 2003). The SNP 1941C>G is a marker for haplotype *IB in block 2 as shown in Table 3. Sai et al. (2004) reported the *IB-dependent decreasing trend of the AUC ratio (SN-38G/SN-38) and an increase in serum total bilirubin levels. Highly differential haplotype distributions in both blocks 1 and 2 observed in different ethnic groups may cause different toxicity profiles during therapy using drugs, such as CPT-11, that are metabolized by UGT1A1.

A novel SNP 686C>A was found in an African-American sample in this study, leading to an amino acid change, P229L, which is different from the known SNP 686C>A (P229Q). Studies using a larger sample size may be necessary to elucidate an accurate frequency. The known SNP 686C>A (P229Q) was reported to be related to Gilbert’s syndrome and to have very low bilirubin glucuronidation activity (Koiwai et al., 1995). However, Jinno et al. (2003a,b) reported that protein expression and mRNA levels of P229Q expressed in COS-1 cells were comparable to those of the wild type, and the decrease in its SN-38 glucuronidation was marginal. Since the association of P229Q with TA7 has been suggested previously (Huang et al., 2000), we have investigated whether the P229Q variant can be mainly attributed to the TA7 variation. On the other hand, the novel variant P229L expressed in COS-1 cells was found to have extremely low SN-38 glucuronidation efficacy in this study (less than 3% of that for the wild type). The low glucuronidation activity of P229L was not caused by its low transcription levels, because a significant difference in the mRNA level was not observed in this study (less than 3% of that for the wild type). The low glucuronidation activity of P229L was not caused by its low transcription levels, because a significant difference in the mRNA level was not observed in this study (less than 3% of that for the wild type). Moreover, a band shifted to a higher molecular weight than the wild-type UGT1A1 was detected by a rabbit anti-human UGT1A antibody in Western blot. Such additional bands were also observed in Western blots of variant UGT1A1 having extremely low enzyme activities, such as F83L and Y486D in our group (Jinno et al., in press). The band shift to a higher molecular weight in SDS-polyacrylamide gel electrophoresis is considered to be the result of products with different post-translational modifications such as phosphorylation or glycosylation, which may occur due to conformational changes caused by amino acid substitution. This may also contribute to the extremely low catalytic activity of the variant enzyme compared with that of the wild type as shown in Table 4. Although the clinical significance of 229L remains to be confirmed, this novel SNP may be a potential cause of diseases such as hyperbilirubinemia or jaundice and may cause severe adverse effects after administration of CPT-11.

Acknowledgments. We thank Professor Ichiro leiri from Tottori University Hospital for extracting the genomic DNA. We also thank Yukult Honsha Co. Ltd. for kindly providing the SN-38 and SN-38G.

References


**Address correspondence to:** Dr. Nahoko Kaniwa, Division of Medicinal Safety Science, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. E-mail: nkaniva@nih.go.jp