RESEARCH ARTICLE

Human Mutation

OFFICIAL JOURNAL
HGVS
HUMAN GENOME VARIATION SOCIETY
www.hgvs.org

Association of Polymorphisms in Four Bilirubin Metabolism Genes with Serum Bilirubin in Three Asian Populations

Rong Lin,[1] Xiaofeng Wang,[1] Yi Wang,[1] Feng Zhang,[1] Ying Wang,[3] Wenqing Fu,[1] Ting Yu,[1] Shilin Li,[1] Momiao Xiong,[4] Wei Huang,[3] and Li Jin[1,2]*

[1]*Ministry of Education Key Laboratory of Contemporary Anthropology and Center for Evolutionary Biology, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai, China*

[2]*Chinese Academy of Sciences-Max Planck Gesellschaft Partner Institute of Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China*

[3]*Chinese National Human Genome Center at Shanghai, Shanghai, China*

[4]*Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas*

Communicated by Daniel W. Nebert

Received 22 April 2008; accepted revised manuscript 25 July 2008.
Published online 25 February 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/humu.20895

ABSTRACT: Numerous studies have shown that the (TA)n repeat polymorphism in the uridine diphosphate glycosyltransferase 1 (*UGT1A1*) gene promoter is associated with hyperbilirubinemia. Several studies also indicated that single nucleotide polymorphism (SNP) rs4148323:G>A at Exon 1 of *UGT1A1* is associated with hyperbilirubinemia. However, it remains unclear what role the polymorphisms play in influencing serum total bilirubin (TBIL) levels in general populations, and whether polymorphisms in other genes involved in the bilirubin metabolism pathway are associated with TBIL levels. The present study addressed these questions by investigating the association of four bilirubin metabolism genes with TBIL levels in three Asian populations: 11 genetic polymorphisms in heme oxygenase-1 (*HMOX1*); biliverdin reductase A (*BLVRA*); solute carrier organic anion transporter family member 1B1 (*SLCO1B1*); and *UGT1A1*. The populations consisted of 502 Kazak herdsmen, 769 Uyghur farmers, and 789 Han farmers, with distinct genetic backgrounds. *UGT1A1* was found to be associated with the $(TA)_7$ allele of the (TA)n repeat polymorphism. We also showed that the A allele of SNP rs4148323:G>A was strongly associated with high TBIL levels in all three populations (each $P<0.005$). Among polymorphisms in other genes, only the (GT)n repeat polymorphism in the *HMOX1* promoter region showed association with TBIL levels in the Uyghur population, but not in the Han and Kazak populations. We also assessed the contributions of (TA)n polymorphism and rs4148323:G>A to phenotypic variations in all three populations. Finally, we observed that significant differences of TBIL levels existed among the three populations; however, this could not be completely explained by the differences at the (TA)n repeat polymorphism and SNP rs4148323:G>A.
Hum Mutat 30, 609–615, 2009.

KEY WORDS: bilirubin; UGT1A1; genetic polymorphism; genetic contribution

Introduction

Bilirubin, a water-insoluble compound, causes hyperbilirubinemia when in excess. Hyperbilirubinemia subsequently increases the risk of schizophrenia [Miyaoka et al., 2000] or leads to acute unconjugated bilirubin encephalopathy, even kernicterus [Brito et al., 2006]. Hence, the genetic mechanism of hyperbilirubinemia has been pursued for decades. However, increasing evidence from experimental data indicated that bilirubin is an effective antioxidant within the normal/mild-increased range. Bilirubin scavenges peroxyl radicals efficiently, suppresses the oxidation of lipids and lipoproteins, especially low density lipoprotein, and thus acts against plaque formation and subsequent atherosclerosis [Mayer, 2000]. Epidemiological studies also showed that serum bilirubin levels are inversely associated with cancer mortality [Temme et al., 2001], coronary artery disease, and mortality from myocardial infarction [Djousse et al., 2001; Mayer, 2000; Vitek et al., 2002]. Therefore, it is of great importance to identify the factors influencing serum bilirubin levels within the normal/mild-increased range.

Serum bilirubin is derived primarily from hemoglobin in aging red blood cells, which is broken down to heme and globin. Heme oxygenases (HMOX) convert heme to biliverdin, which is reduced by biliverdin reductase A (BLVRA; MIM# 109750) to bilirubin. Bilirubin is carried by albumin in the blood and is taken into the liver by solute carrier organic anion transporter family member 1B1 (SLCO1B1; MIM# 604843). Within hepatocytes, the solubility of bilirubin is increased by the addition of one or two molecules of

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Li Jin, PhD, School of Life Sciences, Fudan University, 220 Handan Road, Shanghai 200433, China.
E-mail: ljin007@gmail.com

Contract grant sponsor: Science and Technology Commission of Shanghai Municipality; Grant number: 04dz14003; Contract grant sponsor: National Natural Science Foundation of China; Grant number: 30500291.

glucuronic acid, a process that is catalyzed by uridine diphosphate glycosyltransferase 1 (UGT1A1; MIM# 191740). Bilirubin monoglucuronide and diglucuronide metabolites are then actively transported into the bile.

HMOX are rate-limiting enzymes in heme degradation. Two isoforms of HMOX exists in human: HMOX1 (MIM# 141250) and HMOX2 (MIM# 141251) [Abraham and Kappas, 2005]. *HMOX2* is constitutive, while *HMOX1* can be induced by adverse environmental conditions such as various oxidative agents. Only one population-based study has examined variants of *HMOX* related to serum total bilirubin (TBIL) levels [Endler et al., 2004]. The study showed that carriers of short alleles (<25 GT) of the (GT)n repeat polymorphism in the *HMOX1* promoter region demonstrated higher bilirubin levels compared with noncarriers.

BLVRA reduces biliverdin to bilirubin, binds to HMOX1 protein [Wang and de Montellano, 2003], and regulates HMOX1 enzyme activities [Liu and Ortiz de Montellano, 2000]. Moreover, BLVRA is a regulator for induction of *HMOX1* expression by oxidative stress [Ahmad et al., 2002]. However, no study has pursued the role of genetic variants of *BLVRA* underlying serum bilirubin levels in the literature. The present study hypothesized that single nucleotide polymorphism (SNP) rs699512:A>G, the only common nonsynonymous SNP within *BLVRA* shown in the dbSNP database, might affect TBIL levels.

SLCO1B1 is exclusively expressed at the basolateral membrane of hepatocytes, and transports unconjugated bilirubin from the blood circulation into the liver [Cui et al., 2001]. To date, two population-based studies have examined association of variants of *SLCO1B1* with serum bilirubin levels. The minor allele of SNP rs2306283:C>T and rs4149056:T>C in the *SLCO1B1* gene increases the risk of unconjugated hyperbilirubinemia among Taiwanese adults [Huang et al., 2005]. Taiwanese neonates with the minor allele of rs2306283:C>T were at high risk to develop severe hyperbilirubinemia, whereas those with the minor allele of rs4149056:T>C were not [Huang et al., 2004]. Of note, these two studies were conducted in hyperbilirubinemia patients. It remains unclear what role these variants play in influencing serum bilirubin levels in the general population.

UGT1A1 is another rate-limiting enzyme in bilirubin metabolism. The (TA)n repeat polymorphism in the *UGT1A1* gene promoter has been studied extensively and has been proved to be associated with hyperbilirubinemia. Several studies also indicated that SNP rs4148323:G>A at exon 1 of *UGT1A1* is associated with hyperbilirubinemia. Nevertheless, most of these studies have examined adult or neonatal hyperbilirubinemia and few studies have examined within the normal bilirubin range [Lin et al., 2006; Sai et al., 2004] or in healthy people [Huang et al., 2000; Ki et al., 2003; Zhang et al., 2007]. In addition, a few studies showed that the (TA)n repeat polymorphism of *UGT1A1* was not associated with hyperbilirubinemia [Babaoglu et al., 2006; Muslu et al., 2007]. Therefore, it is necessary to study these associations in other populations and within the normal/mild-increased range.

The present study aimed to investigate, in three general populations with distinctive genetic backgrounds: 1) whether polymorphisms in *HMOX1*, *BLVRA*, and *SLCO1B1*, in addition to *UGT1A1*, are associated with TBIL levels; 2) how genetic variants contribute to the TBIL variation; and 3) whether the genetic difference could account for the phenotypic difference between different populations. To our knowledge, this is the first systematic association study on TBIL that has included all four main genes in bilirubin metabolism and in which TBIL is considered as a quantitative trait. The large sample size is additional strengths of this investigation.

Subjects and Methods

Study Populations

Subjects were recruited from three Asian groups currently residing in the Xinjiang region of China. Subjects were as follows: 502 Kazak herdsmen, 769 Uyghur farmers, and 789 Han farmers. The Kazak and Uyghur are two major ethnic groups distributed across Central Asia and both are admixed groups that are genetically derived from East Asians and Caucasians. These subjects were recruited as natural populations without preselection bias toward their phenotypes. The Kazak herdsmen were enrolled from four villages of Barlikun County, one village of Mulei County, one village of Hutubi County, and one village of Xinyuan County of Xinjiang in May and September of 2003. The Uyghur farmers were recruited from three villages of the rural area of Tulupan District of Xinjiang in March to May of 2005. The Han farmers were recruited from four villages of the rural area of Tulupan District in April of 2006. They migrated to the current location about 50 years ago from Jiangsu Province, China. All participants were invited to have a physical examination and blood tests at a local hospital between 8:00 AM and 1:00 PM after an overnight fast. Participants were asked to avoid alcohol, cigarette smoking, and tea for at least 1 hr before the physical examination. Ethnicity was self-reported. Gallstone and serious liver diseases such as hepatitis were not found in the participants. None of the subjects were related, and all were older than 25 years of age and provided written informed consent. The procedure was approved by a local Ethics Committee. Fasting venous blood samples from the subjects were taken for biochemical analysis. Among them, 4.5% of the participants had TBIL levels >25 μmol/L but all levels were <65 μmol/L, falling into the range of Gilbert's syndrome, which has a prevalence of 3 to 10% in general populations [Bosma et al., 1995]. The characteristics of the three populations are summarized in Table 1.

Genotyping

The information on the 11 polymorphisms genotyped in this study is listed in Table 2. The genotyping method is specified for each polymorphism in the online Supplementary Methods and Supplementary Table S1 (available online at http://www.interscience.wiley.com/jpages/1059-7794/suppmat).

Statistical Analysis

Hardy-Weinberg equilibrium (HWE) for the (GT)n repeat of *HMOX1* and the (TA)n repeat of *UGT1A1* was performed using a homozygosity test [Weir, 1992], a likelihood ratio test [Chakraborty et al., 1991], and an exact test [Guo and Thompson, 1992]. HWE for other loci was examined using a chi-squared test.

Haplotypes and their frequencies were inferred using PHASE version 2.1.1 software (www.stat.washington.edu/stephens/phase/download.html). Based on the inferred haplotype data, we used HaploBlockFinder (http://cgi.uc.edu/cgi-bin/kzhang/haploBlockFinder.cgi) to identify block structures, and to calculate $|D'|$ and r^2 to represent linkage disequilibrium (LD). The Haploview program (www.broad.mit.edu/haploview/haploview) was adopted to yield similar haplotype block structures, $|D'|$, and r^2 compared with those from HaploBlockFinder.

Since the distribution of TBIL levels were skewed, the values were transformed using $\log_{10}$ for all tests of statistical significance. After transformation, the distributions of TBIL levels in the three populations became approximately normal (data not shown). The

Table 1. General Characteristics of Study Participants*

	Han	Uyghur	Kazak
Number of subjects	789	769	502
Men (%)	39.5	36.9	41.5
Age (years)	52.9 ± 12.9	52.8 ± 10.4	48.1 ± 9.7
Body mass index (kg/m^2)	24.6 ± 3.5	26.8 ± 4.5	26.3 ± 4.5
Waist-to-hip ratio	0.874 ± 0.066	0.879 ± 0.086	0.897 ± 0.005 (n = 360)
Total bilirubin (μmol/L)	13.5 ± 5.3	12.9 ± 5.8	14.8 ± 7.3
Fasting glucose (mmol/L)	5.59 ± 1.34	5.76 ± 2.02	5.64 ± 1.18
Total cholesterol (mmol/L)	3.88 ± 0.74	4.52 ± 1.14	5.24 ± 1.07
Triglycerides (mmol/L)	1.12 ± 0.69	1.76 ± 1.30	1.44 ± 0.90

*Values shown are mean ± SD, except for waist-to-hip ratio in the Kazak population, which are mean ± SE.
SD, standard deviation; SE, standard error.

Table 2. Polymorphisms Enrolled in the Study

			Minor allele frequency		
Gene	Polymorphisms	Function	Han	Uyghur	Kazak
HMOX1 (NC_000022.9)	rs2071746:T > A	Promoter region: –413[a]	0.451	0.463	0.504
	(GT)n	Promoter region: –198[a]	0.371[c]	0.462[c]	0.499[c]
	rs2071749:G > A	Intron 3	0.281	0.310	0.329
BLVRA (NC_000007.12)	rs699512:A > G	Exon 2: Thr3Ala	0.329	0.251	0.283
SLCO1B1 (NC_000012.10)	rs4149013:A > G	Promoter region: –12099[b]	0.120	0.100	0.148
	rs17328763:T > C	Promoter region: –11939[b]	0.003	0.066	0.071
	rs4149014:T > G	Promoter region: –11556[b]	0.286	0.153	0.158
	rs2306283:C > T	Exon 5: Asp130Asn	0.229	0.397	0.383
	rs4149056:T > C	Exon 6: Val174Ala	0.112	0.122	0.174
UGT1A1 (NC_000002.10)	(TA)n	Promoter region: –41[b]	0.134	0.256	0.277
	rs4148323:G > A	Exon 1: Gly71Arg	0.211	0.168	0.211

[a]The positions of variations correspond to positions in *HMOX1* with the transcription initiation site set to 1.
[b]The positions of variations correspond to positions in the corresponding cDNA sequence with the first base of the ATG first codon set to 1.
[c]The frequency of Group 2.

differences in serum TBIL levels, classified according to different genotypes, were examined using analysis of variance (ANOVA) and adjusted using analysis of covariance (ANOCVA). For multiple hypothesis testing, a Benjamini and Hochberg false discovery rate (FDR) was applied to the univariable models and a P value ≤ 0.05 (FDR correction) was considered significant. Furthermore, stepwise linear regression analysis using a criterion of $P < 0.05$ was performed to further identify a subset of polymorphisms that significantly explained variance in TBIL levels when adjusted for the effects of other polymorphisms, and to calculate partial R^2 (an increment to R^2 as the polymorphism was added into the linear regression model), which denotes the proportion of the variation in TBIL levels explained by the polymorphism. Candidate covariates in stepwise selection included all the polymorphism effects and the covariates of age, gender, body mass index, waist-to-hip ratio, fasting glucose, total cholesterol, and triglycerides. A general (additive) model was assumed for each polymorphism. The waist-to-hip ratio was excluded in the Kazak population due to incomplete data, while SNP rs17328763:T > C was excluded in the Han population due to the low frequency of the minor allele. All these analyses were performed using the Statistical Package for Social Science (SPSS), version 15.0 (SPSS, Inc., Chicago, IL).

Haplotype-based association analysis, which excluded the subjects with missing genotype data, was conducted with the "testing haplotype effects in association studies" (THESIAS) software (http://genecanvas.ecgene.net/downloads.php?cat_id=1). The haplotype–phenotype association parameters were adjusted for age and gender. Effects associated with rare haplotypes (frequency < 0.02) were not estimated and were set to 0. Because various haplotype reconstruction methods involve different amounts of error [Lamina et al., 2008], we also conducted haplotype analysis using PHASE 2.1.1 combined with ANOVA and ANOCVA to confirm the results.

Results

Genotyping Results

The minor allele frequencies (MAFs) of all polymorphisms are presented in Table 2. As shown, the MAF of SNP rs17328763:T > C of *SLCO1B1* was lower than 0.003 and therefore was excluded in the further analysis involving the Han population. The counts of genotypes at all polymorphisms are presented in Supplementary Table S2. None of the polymorphisms showed statistically significant deviation from HWE.

The (GT)n repeat polymorphism and SNP rs2071746:T > A are two polymorphisms in the promoter region of *HMOX1* and are 154 nucleotides apart from each other. The distributions of the number of (GT)n repeats in the *HMOX1* gene promoter in these three populations were all trimodal, with one peak located at 23 GT repeats and the other two peaks located close together at 30 and 34 GT repeats (Fig. 1). Therefore, the allele type was classified into three classes according to the (GT)n repeat size following Yamada et al. [2000], namely, short alleles (S: < 27 GT), middle alleles (M: 27-32 GT), and long alleles (L: ≥ 33 GT). Similar to Japanese [Ono et al., 2004], A-30 (frequency 30–35%) and T-23

(frequency 16–23%) were the two major haplotypes of rs2071746-(GT)n repeat in every population of this study.

Because there were only 27 L/L carriers in all three populations (Han, 10; Uyghur, 8; and Kazak, 9) and because experimental data showed decreasing *HMOX1* promoter activity with an increasing number of GT repeats [Chen et al., 2002], we merged L/L carriers into M/L carriers for further analysis. Given that the difference of the *HMOX1* expression between L/L and M/L carriers is similar to that between M/L and M/M carriers, we sorted L/L, M/L, and M/M carriers into one group. Furthermore, *HMOX1* expressions and HMOX activities induced by H_2O_2 stimulation were found to be significantly higher in lymphoblastoid cell lines from Japanese subjects with S/S (<27 GT) genotype than those with L/L (≥33 GT) genotype [Hirai et al., 2003]. In addition, Okamoto et al. [2006] showed that a cell with S/M has higher *HMOX1* expression than one with M/M under ultraviolet A irradiation. Therefore, we sorted the subjects carrying S/S or S/M GT repeats into Group 1 and the subjects carrying M/M, M/L, or L/L GT repeats into Group 2. Given that a construct carrying more GT repeats expresses HMOX1 less efficiently than one carrying fewer [Chen et al., 2002], it was reasonable to assume that S/L carriers express HMOX1 less than S/M carriers. We therefore classified the subjects carrying S/L GT repeats into Group 2. Obviously, the *HMOX1* expression of Group 1 is expected to be higher than that of Group 2.

The (TA)n repeat polymorphism is located in the promoter region of *UGT1A1*. In addition to the $(TA)_6$ and $(TA)_7$ alleles of *UGT1A1*, which are common in most populations worldwide, we also found 14 $(TA)_5$ alleles and 3 $(TA)_8$ alleles in the subjects of this study. Because of their low frequencies and reported decreasing promoter activity with an increasing number of TA repeats [Beutler et al., 1998], we merged $(TA)_5$ into $(TA)_6$ and $(TA)_8$ into $(TA)_7$ for further analysis, following Lin et al. [2006].

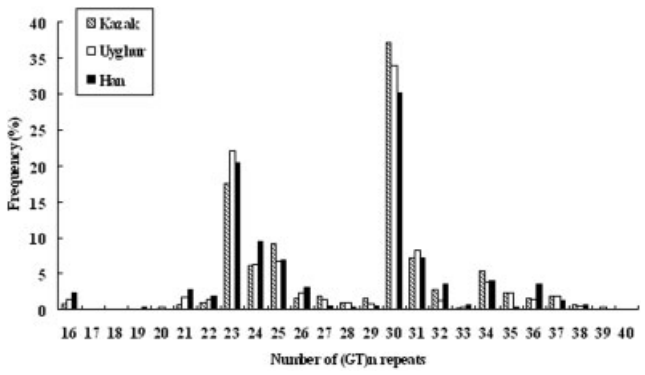

Figure 1. Distribution of the (GT)n repeat of *HMOX1* in this study.

LD Test

Strong pairwise LD was observed between the polymorphisms (each $|D'| > 0.8$) within *HMOX1* and *UGT1A1*, respectively, but with lower r^2 values (each $r^2 < 0.5$) in every population (Supplementary Table S3). Pairwise LD between the polymorphisms within *SLCO1B1* was also low (each $r^2 < 0.5$).

Association with Serum TBIL Levels

Results of association analyses of each of 11 genetic variants with TBIL levels are shown in Supplementary Table S2. As shown, the (TA)n repeat polymorphism and SNP rs4148323:G>A of *UGT1A1* were associated with TBIL levels (age- and gender-adjusted $P = 2.05 \times 10^{-26}$ and $P = 5.21 \times 10^{-16}$, respectively) whereas polymorphisms at the other three genes were not present in all the subjects. The $(TA)_7$ allele of the (TA)n repeat polymorphism and the A allele of rs4148323:G>A were both associated with an increase of TBIL (unadjusted TBIL levels and 95% confidence interval [CI] ranges for specific carriers were as follows: 11.8 [11.6–12.1], 13.0 [12.6–13.4], and 17.7 [16.1–19.5] μmol/L in $(TA)_6/(TA)_6$, $(TA)_6/(TA)_7$, and $(TA)_7/(TA)_7$ carriers, respectively; and 11.4 [11.0–11.8], 12.5 [11.8–13.2], and 14.1 [12.0–16.6] μmol/L in G/G, G/A, and A/A carriers, respectively).

When stratified by ethnic populations, the (TA)n repeat polymorphism and SNP rs4148323:G>A both remained associated with TBIL levels in all three populations (each $P < 0.005$) (Fig. 2; Supplementary Table S2): homozygotes of the minor allele had highest mean TBIL levels, followed by heterozygotes, and then homozygotes of the major allele. Of note, a new association was observed after population stratification. That is, the (GT)n repeat polymorphism of *HMOX1* also showed association with TBIL levels in the Uyghur population (age- and gender-adjusted $P = 0.007$). Group 1 (12.2 μmol/L; 95% CI, 11.8–12.8 μmol/L) was observed to have higher TBIL levels than Group 2 (11.3 μmol/L; 95% CI, 10.8–11.8 μmol/L) in the Uyghur population.

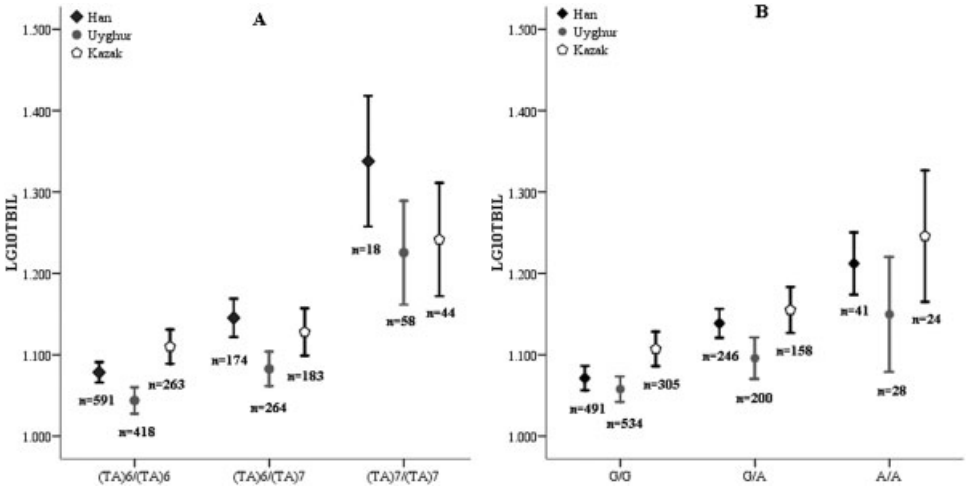

Figure 2. **A**: Mean effects of the *UGT1A1* (TA)n repeat polymorphism on TBIL levels (log_{10} transformed). **B**: Mean effects of SNP rs4148323:G>A on TBIL levels (log_{10} transformed). Unadjusted log_{10} TBIL levels shown are mean ± 95%CI.

In stepwise selection, both the (TA)n repeat polymorphism of *UGT1A1* and SNP rs4148323:G>A remained associated with TBIL levels in each population while the (GT)n repeat polymorphism of *HMOX1* remained associated only in the Uyghur population. The total variation of TBIL levels explained by the (TA)n repeat polymorphism of *UGT1A1* and SNP rs4148323:G>A was 7.1% (stepwise $P = 1.47 \times 10^{-12}$) and 9.8% (stepwise $P = 4.15 \times 10^{-18}$), respectively, in the Han population; 7.0% (stepwise $P = 1.90 \times 10^{-11}$) and 4.4% (stepwise $P = 4.12 \times 10^{-8}$), respectively, in the Uyghur population; and 5.9% (stepwise $P = 4.93 \times 10^{-7}$) and 3.9% (stepwise $P = 6.95 \times 10^{-5}$), respectively, in the Kazak population. The (GT)n repeat polymorphism of *HMOX1* explained 1.1% of the variation of bilirubin in the Uyghur population (stepwise P = 0.006).

In addition, SNP rs699512:A>G at *BLVRA* was found to be associated with TBIL levels in the Han population (P = 0.040), but lost its significance after FDR correction and failed to enter the stepwise linear regression model (Supplementary Table S2).

In the Han population, TBIL levels were significantly lower in women (12.2 μmol/L; 95% CI, 11.8–12.6 μmol/L) than in men (13.2 μmol/L; 95% CI, 12.7–13.8 μmol/L) (P = 0.002) and the gender difference explained 1.1% of the variation of TBIL levels (stepwise P = 0.003). However, the gender differences were not found in the Uyghur and Kazak populations.

To account for the LD between SNP rs4148323:G>A and the (TA)n repeat polymorphism ($|D'| > 0.85$) in every population, a haplotype-based analysis was performed. The frequencies of the $(TA)_7$-A haplotype at the (TA)n repeat polymorphism and rs4148323:G>A were all lower than 0.02 in the three populations, indicating that two alleles associated individually with high TBIL levels rarely existed on the same chromosome. By reference to the most frequent haplotype $(TA)_6$-G, both the $(TA)_6$-A haplotype and the $(TA)_7$-G haplotype were associated with increased TBIL levels (each $P < 10^{-3}$) in all three populations (Supplementary Table S4), suggesting that both the (TA)n repeat polymorphism and SNP rs4148323:G>A contributed to the variation of TBIL levels. These results indicated that the $(TA)_7$ allele of the (TA)n repeat polymorphism and the A allele of SNP rs4148323:G>A are two predictors of high TBIL levels. The haplotype analysis with PHASE 2.1.1 also supported this conclusion: 1) the frequencies of the $(TA)_7$-A haplotype at the (TA)n repeat polymorphism and rs4148323:G>A were all lower than 0.02 in the three populations; and 2) by reference to the most frequent haplotype $(TA)_6$-G, both the $(TA)_6$-A haplotype and the $(TA)_7$-G haplotype were associated with increased TBIL levels (each $P < 10^{-5}$) in all three populations (Supplementary Table S5).

Differences of Serum Bilirubin Levels Among Populations

The mean TBIL levels of the Han participants (12.6 μmol/L; 95% CI, 12.2–12.9 μmol/L) were higher than that of the Uyghur participants (11.8 μmol/L; 95% CI, 11.4–12.1 μmol/L) (P = 0.002), yet lower than that of the Kazak participants (13.4 μmol/L; 95% CI, 12.9–14.0 μmol/L) (P = 0.004) (Table 3). The difference between the Han and Kazak populations remained significant (P = 0.003) after adjustment for rs4148323:G>A, while this difference became nonsignificant (P = 0.300) after adjustment for the (TA)n repeat polymorphism of *UGT1A1*. This indicated that the difference between the Han and Kazak populations could be explained by the distribution of the (TA)n repeat polymorphism of *UGT1A1*. The MAF of rs4148323:G>A in the Han population was almost identical to that in the Kazak population (P = 0.966), while the frequency of the $(TA)_7$ allele (0.134) in the Han population was lower than that in the Kazak population (0.277) ($P = 4.38 \times 10^{-19}$) (Supplementary Table S6). The subjects carrying the $(TA)_7/(TA)_7$ genotype have been shown to have higher TBIL levels compared with those carrying the $(TA)_6/(TA)_6$ genotype in all three populations. Thus, the frequency difference of the $(TA)_7$ allele between the Han and Kazak populations was enough to explain the fact that TBIL levels of the Han population were lower than that of the Kazak population.

However, the differences between the Han and Uyghur populations and the Uyghur and Kazak populations could not be explained by the distribution of the (TA)n repeat polymorphism of *UGT1A1* and rs4148323:G>A alone: after adjusting for the (TA)n repeat polymorphism of *UGT1A1* and rs4148323:G>A, the Uyghur population still showed lower bilirubin levels than the Han ($P = 4.41 \times 10^{-7}$) and Kazak populations ($P = 2.82 \times 10^{-6}$). The MAF of rs4148323:G>A in the Uyghur population (0.168) was lower than that in the Han (0.211) and Kazak (0.211) populations (Supplementary Table S5), which may partially explain the fact that the Uyghur population had lower bilirubin levels than the Han and Kazak populations.

Discussion

The present study indicated that the (TA)n repeat polymorphism and SNP rs4148323:G>A of *UGT1A1* were associated with bilirubin levels in all three Asian populations. It also showed that: 1) only one polymorphism at other three genes (i.e., the (GT)n repeat polymorphism of *HMOX1*) was associated with TBIL levels and that this association was only observed in the Uyghur population; and 2) TBIL levels between the three populations were significantly different and the (TA)n repeat polymorphism and SNP rs4148323:G>A only explained the difference between the Han and Kazak populations.

In the Framingham Heart Study, the $(TA)_7$ allele with the frequency 0.323 is associated with high bilirubin levels, which can explain 18.6% of the total variation of bilirubin [Lin et al., 2006]. Compared with the Framingham Heart study, the R^2 values of the $(TA)_7$ allele were low, accompanied by low $(TA)_7$ allele frequencies in the three Asian populations (0.134 in Han Chinese, 0.256 in Uyghur,

Table 3. TBIL Levels Stratified by Population

Population	TBIL (μmol/L)[a]	Comparing between two populations	P value	P value[b]	P value[c]	P value[d]	P value[e]
Han	12.6 (12.2–12.9)	Han:Uyghur	0.002	0.002	2.05×10^{-7}	0.006	4.41×10^{-7}
Uyghur	11.8 (11.4–12.1)	Han:Kazak	0.004	0.013	0.300	0.003	0.582
Kazak	13.4 (12.9–14.0)	Uyghur:Kazak	2.33×10^{-8}	3.12×10^{-7}	3.47×10^{-7}	4.45×10^{-7}	2.82×10^{-6}

[a]Unadjusted values shown for TBIL levels are geometric means (95% CI).
[b]Adjusted for age and gender.
[c]Adjusted for the (TA)n repeat polymorphism of *UGT1A1*.
[d]Adjusted for SNP rs4148323:G>A.
[e]Adjusted for the (TA)n repeat polymorphism of *UGT1A1* and rs4148323:G>A.

and 0.277 in Kazak), which were consistent with the literature (0.143 in Taiwanese, 0.188 in Malays, and 0.100–0.168 in Japanese) [Akaba et al., 1998; Ando et al., 1998; Balram et al., 2002; Huang et al., 2000; Saeki et al., 2003]. In contrast, the frequency of allele $(TA)_7$ is much higher in Caucasians (0.357–0.415) [Beutler et al., 1998; Monaghan et al., 1997] and Indians (0.351) [Balram et al., 2002].

SNP rs4148323:G>A is another common polymorphism of *UTG1A1* in East Asians (Japanese, MAF = 0.130; Taiwanese, MAF = 0.109) [Akaba et al., 1998; Huang et al., 2000] while it is virtually monomorphic in Caucasians. The MAFs of rs4148323:G>A for the Han, Uyghur, and Kazak populations were 0.211, 0.168, and 0.211, respectively, which explained 9.8%, 4.4%, and 3.9%, respectively, of the total variation of TBIL levels.

However, two studies conducted in East Asians showed that the (TA)n repeat polymorphism and rs4148323:G>A associated with TBIL levels had similar MAF but have higher R^2. A study conducted in 89 Japanese with various cancers displayed that the (TA)n repeat polymorphism with a MAF 0.121 was responsible for the variation of TBIL levels ($R^2 < 20.29\%$) whereas rs4148323 with a MAF 0.146 was not [Sai et al., 2004]. The sample size for this study was relatively small. Thus we can not rule out the possibility of an inadvertent selection bias that cause high R^2 for the (TA)n repeat polymorphism and nonsignificance of rs4148323. Ki et al. [2003] showed that the variability of TBIL levels explained by the (TA)n repeat polymorphism and rs4148323:G>A were 28.1% and 12.9%, respectively, in 324 healthy male Koreans. Compared with the Han population in the present study, the Korean population reported by Ki et al. [2003] had smaller sample size but had similar standard deviation (SD) of TBIL levels (mean ± SD: 16.3 ± 5.9 μmol/L) and similar MAF for the (TA)n repeat polymorphism (0.127) and rs4148323:G>A (0.213), which meant the precision of our measure of TBIL levels was not lower than that of Ki et al. [2003]. When we restricted the study participants to men and even restricted the age range to 38–62 years according to the characteristic of the Korean subjects, the result still showed low R^2 for the (TA)n repeat polymorphism and rs4148323:G>A in all the three populations, except 14.2% for rs4148323:G>A in the Han population. The study of two clinic visits in 84 Utah pedigrees indicated that a major gene explained 27% and 28% of the variance in bilirubin levels at visit 1 and visit 2, respectively, and that 22% of the variance in bilirubin levels could be explained by other genes [Hunt et al., 1996]. A Framingham Heart study, another pedigree-based study, has also showed that the heritability of TBIL levels is estimated to be 49% ± 6% and that *UGT1A1* might be a major gene controlling TBIL levels [Lin et al., 2003]. Only three polymorphisms of *UGT1A1* (i.e., –3279T>G, (TA)n repeat polymorphism, and rs4148323:G>A) could explain 41.0% of the variability of TBIL levels in the Korean population [Ki et al., 2003], which could be attributed to the population having relatively homogeneous genetic and environmental backgrounds.

The effect of the (GT)n repeat polymorphism in the *HMOX1* promoter region on TBIL levels was only observed in the Uyghur population. Mean TBIL levels of Group 1 consisting of the S/S or S/M GT repeats carrier were significantly higher than that of Group 2 consisting of the M/M, S/L, M/L, or L/L GT repeats carrier. This was consistent with the previous functional studies, which suggested that the *HMOX1* expression of Group 1 is higher than that of Group 2. The effect of the (GT)n repeat polymorphism on TBIL levels were different in the three populations, for *HMOX1* is not a major gene controlling TBIL levels, which may make the effect of *HMOX1* easily affected by differences among populations such as distinct genetic backgrounds and living conditions.

Our findings also suggested that the differences of TBIL levels between the Han and Uyghur populations and the Uyghur and Kazak populations remained after adjustment for the (TA)n repeat polymorphism and SNP rs4148323:G>A of the *UGT1A1* gene. African Americans have been reported to have lower bilirubin levels but have higher frequencies of $(TA)_7$ and $(TA)_8$ than Caucasians [Beutler et al., 1998]. One of the reasons for this may be that African Americans have lower hemoglobin levels than Caucasians [Beutler and West, 2005; Perry et al., 1992]. Therefore, we suspected that the Uyghur population had lower hemoglobin levels than the Han and Kazak populations.

The Western Eurasian-specific mtDNA haplogroup frequency was observed in Uygur (42.6%) and Kazak (30.2%) from Xinjiang Province, but not found in Han samples from the same place [Yao et al., 2004], suggesting the admixed feature of Kazak and Uygur with Mongoloid and Caucasian. For instance, the Uygur population of Xinjiang is an admixture population with 60% of ancestry from European and 40% of ancestry from East Asia, and this admixture event is estimated to have taken place about 126 generations or 2,520 years ago [Xu et al., 2008]. All these revealed that each of these three groups is derived from a different descent and their genetic backgrounds are distinct. Different genetic backgrounds may lead to different phenotypic propensities for complex traits. Other gene polymorphisms not investigated in the present study could also account for these differences among the three populations. For example, the *UGT1A1* polymorphisms (–3279T>G [Zhang et al., 2007] and 1941C>G [Saeki et al., 2007]) have been showed to be related to TBIL levels.

Environmental factors may also account for these observed differences between the populations. The Uyghur and Han participants live in the Tulupan area, which is arid and one of the hottest places in the world, while the Kazak participants live in Tianshan Mountain, which is cold and semiarid. Furthermore, the cultures and living habits are diverse in these three groups. For example, the Uyghur and Han group generally live by farming while the Kazak live by herding. Genetic and environmental factors likely play interactive roles. For example, individuals with the $(TA)_7/(TA)_7$ genotype of *UGT1A1* reduce bilirubin levels with increasing intake of cruciferous vegetables, whereas individuals with the $(TA)_6/(TA)_6$ or $(TA)_6/(TA)_7$ genotype do not [Peterson et al., 2005]. Distinct genetic backgrounds and their interactions with different environmental factors may provide information for us to understand the differences among the three populations.

We acknowledge that this study has some limitations. Among common SNPs found in Chinese samples of the HapMap project (HapMap release 22 [Apr. 2007], NCBI Build 36), the SNPs employed in the present study were able to capture 71.4%, 60.0%, 25.7%, and 10.0% of them at $r^2 > 0.80$ in *HMOX1*, *BLVRA*, *SLCO1B1*, and *UGT1A1*, respectively. Therefore, the present study had a limitation in that not all common variations were examined comprehensively at *BLVRA*, *SLCO1B1*, and *UGT1A1*. However, the present study was economical and included almost all the important and known functional common variants in the four bilirubin metabolism genes.

Another limitation is that TBIL levels were measured only once in this cross-sectional study and could possibly have changed if a second measurement were taken. However, on the basis that there is a significant correlation of 0.65 for bilirubin levels and a consistent genetic contribution to bilirubin between the two clinic visits over the 2.5 years in 84 Utah pedigrees [Hunt et al., 1996], we do not anticipate that a second measurement of TBIL levels would be significantly different from the original measurement in our populations and alter the present results.

In summary, the present study demonstrated that the (TA)n repeat polymorphism and SNP rs4148323:G>A of *UGT1A1* are associated with TBIL levels in all three Asian populations and the (GT)n repeat polymorphism in the *HMOX1* promoter region may affect bilirubin levels in the Uyghur population. The (TA)n repeat polymorphism and SNP rs4148323:G>A cannot explain fully the differences of TBIL levels among the three populations. Other polymorphisms that influence TBIL levels need to be further explored to explain these differences.

References

Abraham NG, Kappas A. 2005. Heme oxygenase and the cardiovascular-renal system. Free Radic Biol Med 39:1–25.

Ahmad Z, Salim M, Maines MD. 2002. Human biliverdin reductase is a leucine zipper-like DNA-binding protein and functions in transcriptional activation of heme oxygenase-1 by oxidative stress. J Biol Chem 277:9226–9232.

Akaba K, Kimura T, Sasaki A, Tanabe S, Ikegami T, Hashimoto M, Umeda H, Yoshida H, Umetsu K, Chiba H, Yuasa I, Hayasaka K. 1998. Neonatal hyperbilirubinemia and mutation of the bilirubin uridine diphosphate-glucuronosyltransferase gene: a common missense mutation among Japanese, Koreans and Chinese. Biochem Mol Biol Int 46:21–26.

Ando Y, Chida M, Nakayama K, Saka H, Kamataki T. 1998. The UGT1A1∗28 allele is relatively rare in a Japanese population. Pharmacogenetics 8:357–360.

Babaoglu MO, Yigit S, Aynacioglu AS, Kerb R, Yurdakok M, Bozkurt A. 2006. Neonatal jaundice and bilirubin UDP-glucuronosyl transferase 1A1 gene polymorphism in Turkish patients. Basic Clin Pharmacol Toxicol 98:377–380.

Balram C, Sabapathy K, Fei G, Khoo KS, Lee EJ. 2002. Genetic polymorphisms of UDP-glucuronosyltransferase in Asians: UGT1A1∗28 is a common allele in Indians. Pharmacogenetics 12:81–83.

Beutler E, Gelbart T, Demina A. 1998. Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? Proc Natl Acad Sci USA 95:8170–8174.

Beutler E, West C. 2005. Hematologic differences between African-Americans and whites: the roles of iron deficiency and alpha-thalassemia on hemoglobin levels and mean corpuscular volume. Blood 106:740–745.

Bosma PJ, Chowdhury JR, Bakker C, Gantla S, de Boer A, Oostra BA, Lindhout D, Tytgat GN, Jansen PL, Oude Elferink RP, Chowdhury NR. 1995. The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. N Engl J Med 333:1171–1175.

Brito MA, Silva RF, Brites D. 2006. Bilirubin toxicity to human erythrocytes: a review. Clin Chim Acta 374:46–56.

Chakraborty R, Fornage M, Gueguen R, Boerwinkle E. 1991. Population genetics of hypervariable loci: analysis of PCR-based VNTR polymorphism within a population. EXS 58:127–143.

Chen YH, Lin SJ, Lin MW, Tsai HL, Kuo SS, Chen JW, Charng MJ, Wu TC, Chen LC, Ding YA, Pan WH, Jou YS, Chau LY. 2002. Microsatellite polymorphism in promoter of heme oxygenase-1 gene is associated with susceptibility to coronary artery disease in type 2 diabetic patients. Hum Genet 111:1–8.

Cui Y, Konig J, Leier I, Buchholz U, Keppler D. 2001. Hepatic uptake of bilirubin and its conjugates by the human organic anion transporter SLC21A6. J Biol Chem 276:9626–9630.

Djousse L, Levy D, Cupples LA, Evans JC, D'Agostino RB, Ellison RC. 2001. Total serum bilirubin and risk of cardiovascular disease in the Framingham offspring study. Am J Cardiol 87:1196–1200; A1194; 1197.

Endler G, Exner M, Schillinger M, Marculescu R, Sunder-Plassmann R, Raith M, Jordanova N, Wojta J, Mannhalter C, Wagner OF, Huber K. 2004. A microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with increased bilirubin and HDL levels but not with coronary artery disease. Thromb Haemost 91:155–161.

Guo SW, Thompson EA. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372.

Hirai H, Kubo H, Yamaya M, Nakayama K, Numasaki M, Kobayashi S, Suzuki S, Shibahara S, Sasaki H. 2003. Microsatellite polymorphism in heme oxygenase-1 gene promoter is associated with susceptibility to oxidant-induced apoptosis in lymphoblastoid cell lines. Blood 102:1619–1621.

Huang CS, Luo GA, Huang ML, Yu SC, Yang SS. 2000. Variations of the bilirubin uridine-diphosphoglucuronosyl transferase 1A1 gene in healthy Taiwanese. Pharmacogenetics 10:539–544.

Huang MJ, Kua KE, Teng HC, Tang KS, Weng HW, Huang CS. 2004. Risk factors for severe hyperbilirubinemia in neonates. Pediatr Res 56:682–689.

Huang CS, Huang MJ, Lin MS, Yang SS, Teng HC, Tang KS. 2005. Genetic factors related to unconjugated hyperbilirubinemia amongst adults. Pharmacogenet Genomics 15:43–50.

Hunt SC, Wu LL, Hopkins PN, Williams RR. 1996. Evidence for a major gene elevating serum bilirubin concentration in Utah pedigrees. Arterioscler Thromb Vasc Biol 16:912–917.

Ki CS, Lee KA, Lee SY, Kim HJ, Cho SS, Park JH, Cho S, Sohn KM, Kim JW. 2003. Haplotype structure of the UDP-glucuronosyltransferase 1A1 (UGT1A1) gene and its relationship to serum total bilirubin concentration in a male Korean population. Clin Chem 49:2078–2081.

Lamina C, Bongardt F, Kuchenhoff H, Heid IM. 2008. Haplotype reconstruction error as a classical misclassification problem: introducing sensitivity and specificity as error measures. PLoS ONE 3:e1853.

Lin JP, Cupples LA, Wilson PW, Heard-Costa N, O'Donnell CJ. 2003. Evidence for a gene influencing serum bilirubin on chromosome 2q telomere: a genomewide scan in the Framingham study. Am J Hum Genet 72:1029–1034.

Lin JP, O'Donnell CJ, Schwaiger JP, Cupples LA, Lingenhel A, Hunt SC, Yang S, Kronenberg F. 2006. Association between the UGT1A1∗28 allele, bilirubin levels, and coronary heart disease in the Framingham Heart Study. Circulation 114:1476–1481.

Liu Y, Ortiz de Montellano PR. 2000. Reaction intermediates and single turnover rate constants for the oxidation of heme by human heme oxygenase-1. J Biol Chem 275:5297–5307.

Mayer M. 2000. Association of serum bilirubin concentration with risk of coronary artery disease. Clin Chem 46:1723–1727.

Miyaoka T, Seno H, Itoga M, Iijima M, Inagaki T, Horiguchi J. 2000. Schizophrenia-associated idiopathic unconjugated hyperbilirubinemia (Gilbert's syndrome). J Clin Psychiatry 61:868–871.

Monaghan G, Foster B, Jurima-Romet M, Hume R, Burchell B. 1997. UGT1∗1 genotyping in a Canadian Inuit population. Pharmacogenetics 7:153–156.

Muslu N, Turhan AB, Eskandari G, Atici A, Ozturk OG, Kul S, Atik U. 2007. The frequency of UDP-glucuronosyltransferase 1A1 promoter region (TA)7 polymorphism in newborns and its relation with jaundice. J Trop Pediatr 53:64–68.

Ono K, Goto Y, Takagi S, Baba S, Tago N, Nonogi H, Iwai N. 2004. A promoter variant of the heme oxygenase-1 gene may reduce the incidence of ischemic heart disease in Japanese. Atherosclerosis 173:315–319.

Perry GS, Byers T, Yip R, Margen S. 1992. Iron nutrition does not account for the hemoglobin differences between blacks and whites. J Nutr 122:1417–1424.

Peterson S, Bigler J, Horner NK, Potter JD, Lampe JW. 2005. Cruciferae interact with the UGT1A1∗28 polymorphism to determine serum bilirubin levels in humans. J Nutr 135:1051–1055.

Saeki M, Saito Y, Jinno H, Tohkin M, Kurose K, Kaniwa N, Komamura K, Ueno K, Kamakura S, Kitakaze M, Ozawa S, Sawada J. 2003. Comprehensive UGT1A1 genotyping in a Japanese population by pyrosequencing. Clin Chem 49:1182–1185.

Saeki M, Saito Y, Sai K, Maekawa K, Kaniwa N, Sawada J, Kawamoto M, Saito A, Kamatani N. 2007. A combinatorial haplotype of the UDP-glucuronosyltransferase 1A1 gene (♯60-♯IB) increases total bilirubin concentrations in Japanese volunteers. Clin Chem 53:356–358.

Sai K, Saeki M, Saito Y, Ozawa S, Katori N, Jinno H, Hasegawa R, Kaniwa N, Sawada J, Komamura K, Ueno K, Kamakura S, Kitakaze M, Kitamura Y, Kamatani N, Minami H, Ohtsu A, Shirao K, Yoshida T, Saijo N. 2004. UGT1A1 haplotypes associated with reduced glucuronidation and increased serum bilirubin in irinotecan-administered Japanese patients with cancer. Clin Pharmacol Ther 75:501–515.

Temme EH, Zhang J, Schouten EG, Kesteloot H. 2001. Serum bilirubin and 10-year mortality risk in a Belgian population. Cancer Causes Control 12:887–894.

Vitek L, Jirsa M, Brodanova M, Kalab M, Marecek Z, Danzig V, Novotny L, Kotal P. 2002. Gilbert syndrome and ischemic heart disease: a protective effect of elevated bilirubin levels. Atherosclerosis 160:449–456.

Wang J, de Montellano PR. 2003. The binding sites on human heme oxygenase-1 for cytochrome p450 reductase and biliverdin reductase. J Biol Chem 278:20069–20076.

Weir BS. 1992. Independence of VNTR alleles defined as fixed bins. Genetics 130:873–887.

Xu S, Huang W, Qian J, Jin L. 2008. Analysis of genomic admixture in Uyghur and its implication in mapping strategy. Am J Hum Genet 82:883–894.

Yamada N, Yamaya M, Okinaga S, Nakayama K, Sekizawa K, Shibahara S, Sasaki H. 2000. Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with susceptibility to emphysema. Am J Hum Genet 66:187–195. Erratumn: Am J Hum Genet 2001; 2068:1542.

Yao YG, Kong QP, Wang CY, Zhu CL, Zhang YP. 2004. Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in china. Mol Biol Evol 21:2265–2280.

Zhang A, Xing Q, Qin S, Du J, Wang L, Yu L, Li X, Xu L, Xu M, Feng G, He L. 2007. Intra-ethnic differences in genetic variants of the UGT-glucuronosyltransferase 1A1 gene in Chinese populations. Pharmacogenomics J 7:333–338.