Short Communication

Genetic variants and haplotypes of the *UGT1A9, 1A7* and *1A1* genes in Chinese Han

Xiaoqing Zhang[1,*], Guokun Ao[2,*], Yuewen Wang[1], Wei Yan[1], Min Wang[1], Erfei Chen[1], Fangfang Yang[1] and Jin Yang[1]

[1]*National Engineering Research Center for Miniaturized Detection System, College of Life Sciences, Northwest University, Xi'an, Shaanxi, China.*
[2]*The 309th Hospital of Chinese People's Liberation Army, HaiDian District, Beijing, China.*

Abstract

In this report, we describe combined polymorphisms of the *UGT1A9*, *UGT1A7* and *UGT1A1* genes in 100 unrelated, healthy Chinese Han subjects. The functional regions of these genes were sequenced and comprehensively analyzed for genetic polymorphisms. Thirty variants were detected, including five novel forms. Tentative functional predictions indicated that a Cys → Arg substitution at position 277 in the *UGT1A7* gene could alter the protein conformation and that 12460T > G in the 3'UTR might influence protein translation through specifically expressed miRNAs. *UGT1A9*1b* was a major functional variant in the subjects examined whereas the *1f* allele had a frequency of only 0.5%. A special functional haplotype (GAGAAC) was identified for *UGT1A9*, *1A7* and *1A1*. These findings provide fundamental genetic information that may serve as a basis for larger studies designed to assess the metabolic phenotypes associated with UGT1A polymorphisms. They also provide important data for the implementation of personalized medicine in Chinese Han.

Key words: Chinese Han, genetic polymorphisms, haplotype, UGT1A.

UDP-glucuronosyltransferases (UGTs) are important enzymes in glucuronidation, the major pathway for the removal of endogenous compounds and potentially toxic chemicals ingested in the diet (King *et al.*, 2001; Chu *et al.*, 2009). *UGT1A9* (MIM# 606434), *UGT1A7* (MIM# 606432) and *UGT1A1* (MIM# 191740) are three important members of the *UGT1A* subfamily that form a gene cluster on chromosome 2q37 (Miners *et al.*, 2002; Cecchin *et al.*, 2009). Each isoform comprises a unique alternate exon 1 that is regulated by its own promoter and is linked to the same four downstream exons as the other isoforms (Gong *et al.*, 2001). The enzyme activities of UGT1A9, 1A7 and 1A1 show wide interindividual variability that can lead to the failure of drug therapy, severe drug toxicity or unpredictable adverse side effects. Major mutations in the *UGT1A1* gene can cause a serious metabolic deficiency in bilirubin that leads to a condition known as Gilbert's syndrome. The missense mutation 211G > A (*UGT1A1*6* allele) markedly attenuates (to <10% of normal) or completely abolishes enzyme activity (Tukey and Strassburg, 2000; Verlaan *et al.*, 2004) and reduces activity towards 7-ethyl-10-hydroxy-camptothecin (SN-38) (Huang *et al.*, 2000). *UGT1A7* and *UGT1A9* also contribute to SN-38 glucuronidation (Gagné *et al.*, 2002; Villeneuve *et al.*, 2003). The relative Vmax of the *UGT1A7*3* allele is 5.8-fold lower than the wild-type and confers a slow glucuronidation phenotype (Guillemette *et al.*, 2000).

Genetic polymorphisms that alter enzyme functions are of clinical value in predicting a predisposition to disease and an individual's ability to respond to certain therapeutic agents (Evans and Johnson, 2001). Haplotypes that combine variants of more than one gene can be used to investigate genotype-phenotype associations (Clark, 2004; Kitsios and Zintzaras, 2009). The co-occurrence of single nucleotide polymorphisms (SNPs) or segmental haplotypes with functional changes in *UGT1A9*, *1A7* and *1A1* can result in altered glucuronidation activity (Judson *et al.*, 2000). In this study, we used direct sequencing to screen *UGT1A9*, *1A7* and *1A1* for genetic polymorphisms in their exons, promoters and surrounding introns and inferred the haplotype structures of each region.

The subjects consisted of 100 Chinese Han (50 males and 50 females, 18-40 years old) from Yulin. All of the participants were in a healthy condition based on physical examination and their medical histories. The aim of the investigation and the procedures involved were explained to

the volunteers, all of whom provided written informed consent before participating in the study. The project was reviewed and approved by the Ethics Committee of Northwest University.

Blood samples were obtained from all participants and genomic DNA was isolated from peripheral blood leukocytes by a standard procedure (Okuda *et al.*, 2002). Polymerase chain reaction (PCR) amplifications were done using 14 sets of primers. DNA sequencing reactions were done using ABI PRISM Big Dye Terminator V3.1 cycle sequencing chemistry. The sequences of the PCR and sequencing primers are available upon request. Sequencher 4.10.1 software (Gene Codes Corporation, Ann Arbor, MI, USA) was used for base calling, sequence alignment and polymorphism determination based on the DNA sequences. Novel variants were those not listed in the dbSNP database and having no rs record. Haploview 4.2 software (Massachusetts Institute of Technology, Cambridge, MA, USA) was used to estimate minor allele frequencies (MAF), test for Hardy-Weinberg equilibrium (HWE) and analyze linkage disequilibrium (LD) (Barrett *et al.*, 2005).

Polymorphisms in the transcription factor binding site in the promoter region were analyzed using TFSEARCH (Transcription Factor Binding Sites Search) software (Heinemeyer *et al.*, 1998) and the MAPPER platform (Marinescu *et al.*, 2005). Two prediction programs, PANTHER (Thomas *et al.*, 2003; Thomas and Kejariwal, 2004) and PolyPhen (Ramensky *et al.*, 2002), were used to estimate the possible impact of novel nonsynonymous mutations on protein function by combining structural information and sequence similarity analysis. Swiss-PdbViewer 3.7 (Peitsch *et al.*, 1997) was used to thread a protein primary sequence onto a three-dimentional template. CASTp was used to identify protein pockets and measure the changes in pocket volume produced by novel amino acid substitutions (Binkowski *et al.*, 2003). MicroRNA (miRNA) target prediction was done using PITA (Kertesz *et al.*, 2007), an algorithm that considers site accessibility and hybridization energy to predict potential miRNAs that target *UGT1As* and harbor novel SNPs in their binding sites.

We detected 30 genetic variants, the allele frequencies of which were in Hardy-Weinberg equilibrium (p > 0.05). Eight polymorphisms were found in the com-

Table 1 - Summary of the *UGT1A9, 7* and *1* polymorphisms detected in Chinese Han.

Location	Nucleotide change	Amino acid change	Number of subjects			Minor allele frequency	HWE *P* value	Remark
			W/W	W/m	M/m			
1A9 Promoter	-2189T > C		97	3	0	0.015	1.000	novel
	-1888T > G		91	9	0	0.045	1.000	rs6731242
	-1819T > C		37	43	20	0.415	0.322	rs13418420
	-441C > T		96	4	0	0.020	1.000	rs2741045
	-332T > C		88	11	1	0.065	0.681	rs2741046
	-118delT		35	47	18	0.415	0.868	rs67695772
	-40C > G		99	1	0	0.005	1.000	novel
1A7 Promoter	-543A > G		54	37	9	0.275	0.593	rs4530361
	-341C > T		75	22	3	0.140	0.576	rs28946877
	-57T > G		60	32	8	0.225	0.419	rs7586110
1A7 Exon1	33C > A	Pro11Pro	58	32	10	0.260	0.143	rs7577677
	387T > G	Asn129Lys	37	44	19	0.414	0.495	rs17868323
	391C > A	Arg131Arg	37	44	19	0.414	0.495	rs17863778
	392G > A	Arg131Gln	37	44	19	0.414	0.495	rs17868324
	622T > C	Trp208Arg	54	36	10	0.280	0.378	rs11692021
	756G > A	Leu252Leu	76	22	2	0.130	1.000	rs17864686
	829T > C	Cys277Arg	99	1	0	0.005	1.000	novel
1A1 Promoter	-364C > T		82	15	3	0.105	0.139	rs887829
	-64G > C		95	5	0	0.025	1.000	rs873478
	-52insTA		81	17	2	0.105	0.568	rs3064744
1A1 Exon 1	211G > A	Gly71Arg	54	39	7	0.265	1.000	rs4148323
	686C > A	Pro229Gln	97	3	0	0.015	1.000	rs35350960
1A Exon 2	6833G > A	Lys317Lys	98	2	0	0.010	1.000	novel
1A Intron 2	6893T > C		91	9	0	0.045	1.000	rs4148327
1A Exon 4	7939C > T	Pro364Leu	95	5	0	0.025	1.000	rs34946978
1A Exon 5	12126T > G	Tyr486Asp	97	3	0	0.015	1.000	rs34993780
1A 3'UTR	12460T > G		99	1	0	0.005	1.000	novel
	12483T > C		2	25	73	0.145	1.000	rs10929303
	12611G > C		2	21	77	0.130	1.000	rs1042640
	12712G > C		2	23	75	0.135	1.000	rs8330

W – wild-type, M – mutant, HWE – Hardy-Weinberg equilibrium.

mon exons region, including five rare sites: 6833G > A, 6893T > C, 7939C > T (Pro364Leu), 12126T > G (Tyr486Asp) and 12460T > G. The novel nucleic acid replacement of G to A at nucleotide 6833 was a synonymous mutation (Lys317Lys) in exon 2. In the specific promoter and exon 1 regions we detected five well-known SNPs for the *UGT1A1* gene, seven promoter variants for the *UGT1A9* gene including two novel polymorphisms (-2189T > C, 1.5%, and -40 > G, 0.5%), and nine common variants and one novel rare nonsynonymous mutation (829T > C, Cys277Arg) for the *UGT1A7* gene. Table 1 summarizes the data for these polymorphisms.

Fourteen alleles were identified across the sequencing regions of the *UGT1A9*, *1A7* and *1A1* genes. There were six well-known alleles located in the promoters that influenced enzyme activity by affecting pre-transcriptional control. Among these alleles, a common SNP rs67695772 (*UGT1A9*1b*) that leads to a deletion of T at -118 position occurred at a high frequency (41.5%), with a genotype distribution (47 heterozygous and 18 homozygous subjects) similar to that of *UGT1A7*2*. In contrast, the allele *UGT1A9*1f*, which combined *1b* with three other sites (-1819T > C, -441C > T and -332T > C) in the promoter region of the *UGT1A9* gene, showed a very low frequency (0.5%). In the coding regions, four variants (*UGT1A7*2*, *3*, *4* and *11*) with a frequency of 25% were found in exon 1 of the *UGT1A7* gene while another four variants (*UGT1A1*6*, *7*, *27* and *63*) were located in the *UGT1A1* gene; the latter variants had relatively low frequencies and none were homozygous. Tables 1 and 2 summarize these findings.

The pairwise LD for polymorphisms in the *UGT1A9*, *1A7* and *1A1* genes was expressed as r^2 on a graded red scale. Among the variants listed in Table 1, only common SNPs with frequencies >10% were tested for pairwise LD. The LD map is shown in Figure S1a in which increasing color intensity indicates greater linkage. In view of the potential of rare mutations to alter enzyme function all of the variants were included in the LD analyses. As shown in Figure S1b, two main linkage blocks were observed across the locus. Block 1 located in the promoter and exon 1 of the *UGT1A7* gene included three complete LD variants (387T > G, 391C > A and 392G > A) that showed the greatest linkage (densest red color). In addition, three polymorphisms in block 2 also had high LD in the 3'-untranslated region (UTR) of the *UGT1A1* gene.

Haplotype analysis using an accelerated EM algorithm identified two major haplotypes, TCTCGT in block 1 and CCC in block 2 that accounted for 57% and 85.5% of the total haplotype diversity, respectively; four rare haplotypes had a frequency of <5%. The level of recombination between the two blocks was predicted to have a multiallelic D' value of 0.35. The significant LD seen here among several *UGT1A* SNPs suggests that it would be possible to use only a small number of highly informative SNP markers as

Table 2 - Ethnic differences in *UGT1A* allelic frequencies.

Ethnicity	N	UGT1A1								UGT1A7					UGT1A9				Reference
		*6	*7	*27	*28	*63	*80	*81	*2	*3	*4	*11	*1b	*1d	*1f	
Han	100	26.5	1.5	1.5	10.5	2.5	10.5	2.5	41.4	27.0	28.0	41.4	41.5	2.0	0.5	This study
Chinese	280	9.9			15.3				37.0	26.3	25.2					Chen *et al.* (2006)
Dong	273	12.5		2.2	7.8											Zhang *et al.* (2007)
She	264			0.9												Zhang *et al.* (2007)
Taiwanese	441	21.3		5.0	28.6				38.8	26.8						Tang *et al.* (2005)
Singaporean	241	7.1	0.2	1.7	18.7	0.6										Zhou *et al.* (2009)
Japanese	301	15.3	0.2	1.0	13.0	0.5					21.4	35.0	66.8			Saeki *et al.* (2006)
Korean	50	11.0	1.0	2.0	14.0		14.0	3.0			21.0	39.0	61.0			Yea *et al.* (2008)
Asian	46	14.0		5.0	18.0				51.0	21.7	22.0	51.0	49.0			Thomas *et al.* (2006)
Caucasian	92	0.0		0.0	28.6				41.0	35.9	36.0	41.0	59.0			Thomas *et al.* (2006)

N – number of subjects in each study.

htSNP for genotyping in association studies. Table 3 summarizes the haplotype frequency and htSNP for each LD block.

The TFSEARCH and MAPPER platforms were used to predict the effects that the two novel polymorphisms in the promoter of the *UGT1A9* gene would have on the transcription factor binding site. Neither the wild-type nor mutant of the -40C > G variant was predicted to bind to any transcription factor. For the -2189T > C variant, analysis of the E-values revealed some differences between the wild-type and mutant, but neither had a consistent effect on the transcription factor binding site. Further studies are required to determine whether the -2189T > C variant has a role in transcription factor regulation.

Ten miRNAs containing the novel polymorphism 12460T>G in the 3'UTR met the threshold of interaction energy for miRNA target pairing. Of these, has-miR-181a and has-miR-26a lost their important binding site with a G substitution and showed a significant difference (p < 0.005) in expression between colon tumor and normal samples (Cummins *et al.*, 2006). The enhanced or reduced miRNA binding caused by SNP variation may have consequences for the corresponding downregulation or upregulation of mRNA stability and translation (Mishra *et al.*, 2007; Chen

et al., 2008). Table S1 shows the miRNA sequence and binding capacity for this variant.

PANTHER subPSEC scores of less than -3.0 (Pdeleterious >0.5) and a PolyPhen PSIC score difference >2.0 indicate that mutations may probably damage protein function. Based on the PANTHER prediction, which yielded a low subPSEC score (-6.08576) and significantly high Pdeleterious value (0.9563), the C277R variant has a good probability of impairing protein function. PolyPhen analysis showed that this variant also tended to enlarge the protein cavity (PSIC score difference = 3.874), yet a further indication of its potential to impair protein function.

CASTp analysis provided a better understanding of the fatal transformation in cavity structure caused by the nonsynonymous SNP (Figure 1A,B). The wild-type cavity (Pocket ID = 40), which has a surface area of 60.9 $Å^2$ and a volume of 42 $Å^3$, is composed of atoms from five residues, namely, Cys277, Gly280, Leu357, His360 and Met362 (Figure 1A). Surprisingly, the replacement of Arg277 with Cys increases the surface area by 55.3 $Å^2$ and results in a greater than two-fold expansion in volume (Pocket ID = 60, surface area = 116.2 $Å^2$, volume = 98.4 $Å^3$). The dramatic change in cavity dimensions in the mutant reflects the recruitment of atoms from an additional seven residues,
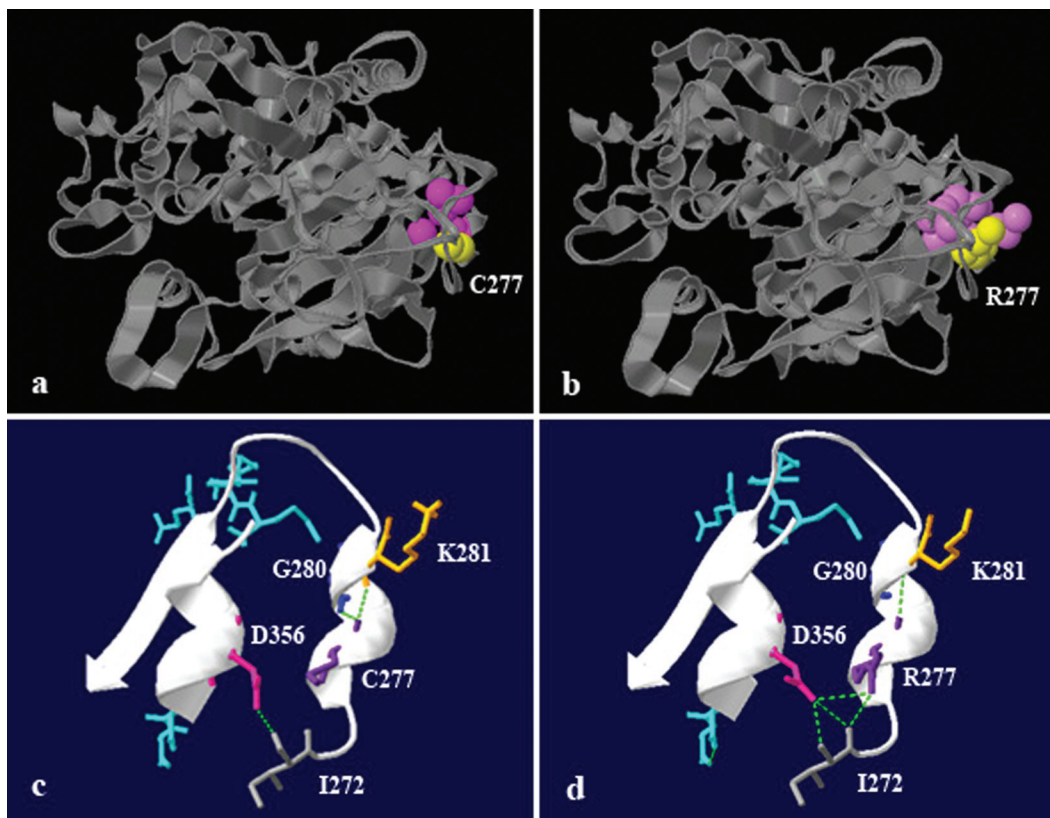

Figure 1 - Functional prediction for the novel nonsynonymous mutation 829T > C in the *UGT1A7* gene. (A, B) The Cys277Ary substitution enlarges the cavity configuration as shown by red balls. The pocket ID 40 in the wide-type (A) is formed by atoms from five residues while the pocket ID 60 in the mutant (B) is formed by atoms from seven residues. (C, D) The Cys277Arg substitution alters hydrogen bonds with the surrounding residues as shown by the green dotted line. Cys277 forms hydrogen bonds with Gly280 and Lys281 (C) whereas Arg277 regenerates hydrogen bonds with Lys281, Ile272 and Asp356 (D).

Table 3 - Haplotype analysis of each LD block in 100 Chinese Han subjects.

Block	Markers	Haplotype	Population frequency	htSNPs
a	-57T > G, 33C > A, 387T > G, 391C > A, 392G > A, 622T > C	TCTCGT	0.570	-57T > G, 33C > A, 387T > G, 622T > C
		GAGAAC	0.218	
		TCGAAT	0.134	
		TAGAAC	0.037	
		GCTCGC	0.013	
		GCGAAT	0.011	
b	12483T > C,12611G > C, 12712G > C	CCC	0.855	12483T > C, 12611G > C
		TGG	0.130	
		TCC	0.010	

htSNPs – haplotype tag SNPs.

namely, Arg277, Gly280, Lys281, Asp356, Leu357, His360 and Met362 (Figure 1B).

The Cys277Arg substitution disrupts a hydrogen bond with Gly280 on the same helix structure but creates two hydrogen bonds with Ile272 and Asp356, and one between them (Figure 1C, D). The three-dimensional structure of the region shown in Figure 1 indicates that the change in hydrogen bonding can explain the expansion of the pocket. These findings suggest that Cys277, together with Ile272 and Asp356, may act as anchors between two helices to regulate the pocket size during interaction with various drugs. The presence of a highly conserved residue in an important domain, together with the striking alteration in cavity configuration caused by this single amino acid substitution, suggests that the new variant could produce a malfunctioning protein.

As shown here, *UGT1A7*11* was a common coding variant (41.4%) in Chinese Han, in agreement with previous investigations of Asian populations (35% in 301 Japanese, 39% in 50 Koreans) (Saeki *et al.*, 2006; Yea *et al.*, 2008). In contrast, the frequencies of other variants differ considerably among ethnic populations. In Asians, the frequencies of *UGT1A1*6* and *27* were reported to range from 7.1% to 21.3% and 0.9% to 5%, respectively (Chen *et al.*, 2006; Zhou *et al.*, 2009), whereas these variants were not observed in Caucasians (Thomas *et al.*, 2006; Borucki *et al.*, 2009). Among Taiwanese, the frequencies of alleles *UGT1A1*2*, *27* and *28* were 21.3%, 5.0% and 28.6%, respectively (Tang *et al.*, 2005), compared with the corresponding value of 18.5%, 1.5% and 10.5% observed here. Similar interethnic differences in frequency exist among Dong, She and Han populations (Zhang *et al.*, 2007). These differences may partly reflect the fact that some rare variants are unique to certain ethnic groups, but may also be related to the small sample sizes used. In this study, we used 100 subjects and this number may have been too small for adequate LD analysis, particularly since D' values tend to be higher when small sample sizes or rare alleles are examined. A larger sample size is therefore required to obtain more precise information.

Although we have predicted the potential significance of novel variants in Chinese Han, additional studies *in vivo* and *in vitro* are required to determine whether these variants cause functional changes in UGT1A. Overall, the special haplotype patterns and htSNPs described here should facilitate genotype-phenotype studies in larger populations.

Acknowledgments

References

Barrett JC, Fry B, Maller J and Daly MJ (2005) Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-265.

Binkowski TA, Naghibzadeh S and Liang J (2003) CASTp: Computed atlas of surface topography of proteins. Nucleic Acids Res 31:3352-3355.

Borucki K, Weikert C, Fisher E, Jakubiczka S, Luley C, Westphal S and Dierkes J (2009) Haplotypes in the UGT1A1 gene and their role as genetic determinants of bilirubin concentration in healthy German volunteers. Clin Biochem 42:1635-1641.

Cecchin E, Innocenti F, D'Andrea M, Corona G, Mattia ED, Biason P, Buonadonna A and Toffoli G (2009) Predictive role of the UGT1A1, UGT1A7 and UGT1A9 genetic variants and their haplotypes on the outcome of metastatic colorectal cancer patients treated with fluorouracil, leucovorin and irinotecan. J Clin Oncol 27:2457-2465.

Chen K, Jin MJ, Zhu YM, Jiang QT, Yu WP, Ma XY and Yao KY (2006) Genetic polymorphisms of the uridine diphosphate glucuronosyltransferase 1A7 and colorectal cancer risk in relation to cigarette smoking and alcohol drinking in a Chinese population. J Gastroenterol Hepatol 21:1036-1041.

Chen K, Song F, Calin GA, Wei Q, Hao X and Zhang W (2008) Polymorphisms in microRNA targets: A gold mine for molecular epidemiology. Carcinogenesis 29:1306-1311.

Chu XY, Liang YX, Cai XX, Cuevas-Licea K, Rippley RK, Kassahun K, Shou M, Braun MP, Doss GA, Anari MR, *et al.* (2009) Metabolism and renal elimination of gaboxadol in

humans: Role of UDP-glucuronosyltransferases and transporters. Pharm Res 26:459-468.

Clark AG (2004) The role of haplotypes in candidate gene studies. Genet Epidemiol 27:321-333.

Cummins JM, He YP, Leary RJ, Pagliarini R, Diaz LA, Barad O, Sjoblom T, Bentwich Z, Szafranska AE, Labourier E, *et al.* (2006) The colorectal microRNAome. Proc Natl Acad Sci USA 103:3687-3692.

Evans WE and Johnson JA (2001) Pharmacogenomics: The inherited basis for interindividual differences in drug response. Annu Rev Genomics Hum Genet 2:9-39.

Gagné JF, Montminy V, Belanger P, Journault K, Gaucher G and Guillemette C (2002) Common human UGT1A polymorphisms and the altered metabolism of irinotecan active metabolite 7-ethyl-10-hydroxycamptothecin (SN-38). Mol Pharmacol 62:608-617.

Gong QH, Cho JW, Huang T, Potter C, Gholami N, Basu NK, Kubota S, Carvalho S, Pennington MW, Owens IS, *et al.* (2001) Thirteen UDP-glucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. Pharmacogenetics 11:357-368.

Guillemette C, Ritter JK, Auyeung DJ, Kessler FK and Housman DE (2000) Structural heterogeneity at the UDP-glucuronosyltransferase 1 locus: Functional consequences of three novel missense mutations in the human UGT1A7 gene. Pharmacogenetics 10:629-644.

Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, *et al.* (1998) Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucleic Acids Res 26:364-370.

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

Judson R, Stephens JC and Windemuth A (2000) The predictive power of haplotypes in clinical response. Pharmacogenomics 1:15-26.

Kertesz M, Iovino N, Unnerstall U, Gaul U and Segal E (2007) The role of site accessibility in microRNA target recognition. Nat Genet 39:1278-1284.

King CD, Rios GR, Green MD and Tephly TR (2001) UDP-glucuronosyltransferases. Curr Drug Metab 1:143-161.

Kitsios GD and Zintzaras E (2009) Genomic convergence of genome-wide investigations for complex traits. Ann Hum Genet 73:514-519.

Marinescu VD, Kohane IS and Riva A (2005) MAPPER: A search engine for the computational identification of putative transcription factor binding sites in multiple genomes. BMC Bioinformatics 6:e79.

Miners JO, McKinnon RA and Mackenzie PI (2002) Genetic polymorphisms of UDP-glucuronosyltransferases and their functional significance. Toxicology 181:453-456.

Mishra PJ, Humeniuk R, Longo-Sorbello GS, Banerjee D and Bertino JR (2007) A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. Proc Natl Acad Sci USA 104:13513-13518.

Okuda T, Fujioka Y, Kamide Y, Kawano Y, Goto Y, Yoshimasa Y, Tomoike H, Iwai N, Hanai S and Miyata T (2002) Verification of 525 coding SNPs in 179 hypertension candidate genes in the Japanese population: Identification of 159 SNPs in 93 genes. J Hum Genet 47:387-394.

Peitsch MC, Wilkins MR, Tonella L, Sanchez JC, Appel RD and Hochstrasser DF (1997) Large scale protein modelling and integration with the SWISS-PROT and SWISS-2DPAGE databases: The example of *Escherichia coli*. Electrophoresis 18:498-501.

Ramensky V, Bork P and Sunyaev S (2002) Human non-synonymous SNPs: Server and survey. Nucleic Acids Res 30:3894-3900.

Saeki M, Saito Y, Jinno H, Sai K, Ozawa S, Kurose K, Kaniwa N, Komamura K, Kotake T, Morishita H, *et al.* (2006) Haplotype structures of the UGT1A gene complex in a Japanese population. Pharmacogenomics J 6:63-75.

Tang KS, Chiu HF, Chen HH, Eng HL, Tsai CJ, Teng HC and Huang CS (2005) Link between colorectal cancer and polymorphisms in the uridine-diphosphoglucuronosyl transferase 1A7 and 1A1 genes. World J Gastroenterol 11:3250-3254.

Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A and Narechania A (2003) PANTHER: A library of protein families and subfamilies indexed by function. Genome Res 13:2129-2141.

Thomas PD and Kejariwal A (2004) Coding single-nucleotide polymorphisms associated with complex *vs.* Mendelian disease: Evolutionary evidence for differences in molecular effects. Proc Natl Acad Sci USA 101:15398-15403.

Thomas SS, Li SS, Lampe JW, Potter JD and Bigler J (2006) Genetic variability, haplotypes, and htSNPs for exons 1 at the human UGT1A locus. Hum Mutat 27:e717.

Tukey RH and Strassburg CP (2000) Human UDP-glucuronosyltransferases: Metabolism, expression, and disease. Annu Rev Pharmacol Toxicol 40:581-616.

Verlaan M, te Morsche RHM, Pap A, Laheij RJF, Jansen JBMJ, Peters WHM and Drenth JPH (2004) Functional polymorphisms of UDP-glucuronosyltransferases 1A1, 1A6, and 1A8 are not involved in chronic pancreatitis. Pharmacogenetics 14:351-357.

Villeneuve L, Girard H, Fortier LC, Gagné JF and Guillemette C (2003) Novel functional polymorphisms in the UGT1A7 and UGT1A9 glucuronidating enzymes in Caucasian and African American subjects and their impact on the metabolism of 7-ethyl-10-hydroxycamptothecin and flavopiridol anticancer drugs. J Pharmacol Exp Ther 307:117-128.

Yea SS, Lee SS, Kim WY, Liu KH, Kim H, Shon JH, Cha IJ and Shin JG (2008) Genetic variations and haplotypes of UDP-glucuronosyltransferase 1A locus in a Korean population. Ther Drug Monit 30:23-34.

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

Zhou YY, Lee LY, Ng SY, Hia CPP, Low KT, Chong YS and Goh DLM (2009) UGT1A1 haplotype mutation among Asians in Singapore. Neonatology 96:150-155.

Internet Resources

UGT Alleles Nomenclature, http://galien.pha.ulaval.ca/labocg/alleles/alleles.html (October 26, 2011).

PANTHER, http://www.pantherdb.org/ (October 28, 2011).

PolyPhen, http://genetics.bwh.harvard.edu/pph/ (October 28, 2011).

PDB, http://www.rcsb.org/pdb/home/home.do (November 5, 2011).

CASTp, http://sts-fw.bioengr.uic.edu/castp/calculation.php (November 5, 2011).

miRBase, http://www.mirbase.org/ (October 28, 2011).

Supplementary Material

The following online material is available for this article:

Table S1 - Prediction of novel SNP that can influence miRNA targeting.

Figure S1 - Haploview analysis of the *UGT1A9*, *1A7* and *1A1* polymorphisms in this study.

This material is available as part of the online article from http://www.scielo.br/gmb.

Associate Editor: Mara H. Hutz

Table S1 - Prediction of novel SNPs that may influence the targeting of miRNAs.

miRNA identification	miRNA sequence	Wild-type score ($\Delta\Delta G$: kcal/mol)	Mutation score ($\Delta\Delta G$: kcal/mol)	miRNA expression in colon		
				Tumor	Normal	p
has-miR-181a	AACAUUCAAcgcugucggugagu	-6.06	-	5	62	0.001832
has-miR-181c	AACAUUCAAccugucggugagu	-4.16	-	3	5	
has-miR-26a	UUCAAGUAAuccaggauaggcu	-2.81	-	125	868	0
has-miR-653	GUGUUGAAAcaaucucuacug	-1.08	-			
has-miR-203	GUGAAAUGUuuaggaccacuag	-0.21	-	6	18	
has-miR-190b	UGAUAUGUUugauauuggguu	-0.18	-	0	4	
has-miR-181b	AACAUUCAUugcugucggugggu	-	-6.96	2	6	
has-miR-181d	AACAUUCAUuguugucgguggggu	-	-3.25	1	1	

miRNA seed sequences are highlighted in uppercase and the complementary SNP sites in the seed sequences are underlined.
miRNAs expression data in colon (tumor/normal samples) was cited from Cummins *et al* (2006) the colorectal microRNAome.
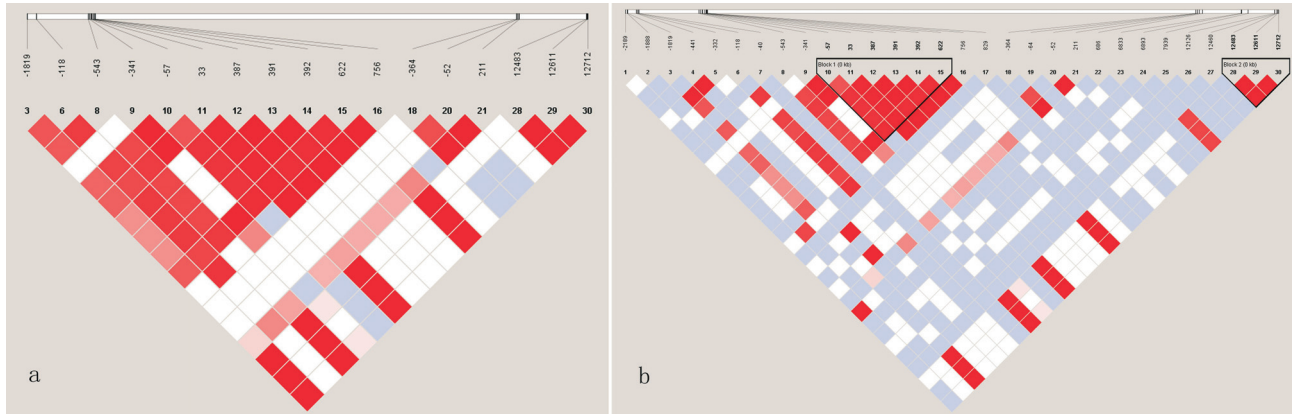Proc Natl Acad Sci USA 103:3687-3692.

Figure S1 - Haploview analysis of the *UGT1A9*, *1A7* and *1A1* polymorphisms in this study. (A) Pairwise LD map for common *UGT1A* SNPs. A denser color indicates greater linkage. (B) LD block for all *UGT1A* SNPs identified in Chinese Han subjects. An area corresponding to a haplotype block is displayed in frame.