Human Molecular Genetics, 2007, Vol. 16, No. 5 ***463–470***
doi:10.1093/hmg/ddl468
Advance Access published on December 22, 2006

Human *SULT1A1* gene: copy number differences and functional implications

Scott J. Hebbring[1], Araba A. Adjei[2], Janel L. Baer[2], Gregory D. Jenkins[3], Jianping Zhang[2], Julie M. Cunningham[1], Daniel J. Schaid[3], Richard M. Weinshilboum[2] and Stephen N. Thibodeau[1,*]

[1]Department of Laboratory Medicine and Pathology, [2]Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics and [3]Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, MN 55905, USA

Received October 13, 2006; Revised and Accepted December 11, 2006

SULT1A1, which catalyzes the sulfate conjugation of a wide variety of natural and synthetic compounds, is genetically polymorphic. Biochemical and pharmacogenetic studies have demonstrated that individual variation in the level of enzyme activity is inherited. Common single-nucleotide polymorphisms (SNPs) located in the open reading frame and in the 5′-flanking region (5′-FR) may account for a portion of this individual variation. In this study, we demonstrate the presence of *SULT1A1* gene deletions and duplications, representing an additional source of variability in the metabolic activity of this enzyme. A quantitative multiplex PCR assay was used to measure the extent of copy number differences and the frequency of these events in different populations. An analysis of DNA from 362 Caucasian-American and 99 African-American showed the presence of 1 to ~5 copies of *SULT1A1* in individual samples: 5% of Caucasian subjects contained a single copy of the gene and 26% had three or more copies, while 63% of African-American subjects had three or more copies. Analysis of the genomic region surrounding the *SULT1A1* gene in three separate cases with a deletion demonstrated that the entire *SULT1A1* gene was affected. Reporter assays, constructed for each of the various 5′-FR SNP haplotypes, suggest that these may also play a role in SULT1A1 activity. However, the variability in the level of enzyme activity among 23 human platelet and 267 human liver samples was best explained by gene copy number differences when all sources of genetic variability were considered ($P < 0.0001$). Overall, these observations have obvious implications for the effectiveness of SULT1A1 as a drug and hormone metabolizing enzyme and its potential role as a risk factor for disease.

INTRODUCTION

Human cytosolic sulfotransferase (SULT) enzymes catalyze the sulfate conjugation of many drugs, other xenobiotics and hormones, especially steroid hormones (1). There are 12 known human SULT isoforms (2–5). *SULT1A1* is one of four *SULT1A* genes that maps to the short arm of chromosome 16 and which encodes proteins that share more than 93% amino acid sequence identity (4,6,7). This area of chromosome 16 has been shown to contain many repetitive sequences and to be highly duplicated (8), as demonstrated by the recent discovery of a large duplication containing *SULT1A3/SULT1A4* (4).

SULT1A1 is ubiquitously expressed in human tissues and is involved in the metabolism, bioactivation and detoxification of environmental and dietary procarcinogens as well as a number of steroid hormones, including estrogens (2,3,9). Phenotypic studies have demonstrated large variations in the levels of human platelet SULT1A1 activity and thermal stability, variations that are largely due to inheritance (10,11). When the *SULT1A1* open reading frame was sequenced using DNA from 33 subjects selected for extreme phenotypes for platelet

*To whom correspondence should be addressed at: Mayo Clinic College of Medicine, 200 First Street SW, 920 Hilton Building, Rochester, MN 55905, USA. Tel: +1 5072849185; Fax: +1 5072840670; Email: sthibodeau@mayo.edu

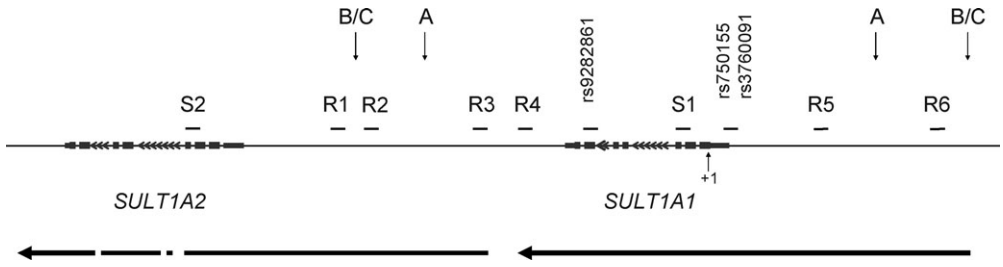

Figure 1. Schematic diagram of gene structure of *SULT1A1* (NM_177534) and *SULT1A2* (NM_177528). Using +1 of A at the ATG start site as the reference point, primer positions for the various assays are as follows: Copy number control (S2, +13448 to +13242), Map set 1 (R1, +10739 to +10595), Map set 2 (R2, +10196 to + 9965), Map set 3 (R3, +6268 to +6125), Map set 4 (R4, +4682 to +4506), Arg213His rs9282861 (+2829 to +2497), Copy number (S1, +524 to +314), (−396) rs750155 and (−624) rs3760091 (−379 to −772), Map set 5 (R5, −2255 to −2399) and Map set 6 (R6, −5499 to −5677). Arrows below represent regions of high homology between *SULT1A1* and *SULT1A2*, and arrows above represent breakpoints for subjects A, B, C.

SULT1A1, a non-synonymous coding single-nucleotide polymorphism (SNP), G638A (Arg213His) or *SULT1A1*2*, was associated with decreased levels of both enzyme activity and thermal stability (12,13). The *SULT1A1*2* polymorphism has been reported to have frequencies of 0.33, 0.29 and 0.80 in Caucasian-American, Han Chinese, and African-American subjects, respectively (14). Subsequently, two additional polymorphisms in the 5′-FR of *SULT1A1* were also reported to contribute to variation in SULT1A1 enzyme activity (15,16). These 5′-FR SNPs, however, were found to be in linkage disequilibrium with the *SULT1A1*2* ORF polymorphism (15,16).

Given the importance of this enzyme in a variety of metabolic processes, a number of studies have examined the role of these polymorphic variants in cancer risk. However, case–control studies testing for the possible association of the *SULT1A1*2* polymorphism with risk for breast (17–19), bladder cancer (20), prostate cancer (20,21) and colorectal, esophageal or lung cancer (3) have produced inconsistent results. These inconsistencies may result from the influence of other SNPs, such as those described in the 5′-FR or as yet undiscovered polymorphisms.

Variability in gene copy number is a common phenomenon throughout the human genome (22–24) and has been observed for a number of genes encoding drug metabolizing enzymes, including the cytochrome P-450 isoform CYP2D6 (25,26). Therefore, after obtaining preliminary indications that *SULT1A1* might display variation in copy number, we set out to test that hypothesis, as well as to explore the possible functional implications of these copy number differences. In the present study, we report on the occurrence of common copy number variations in the human *SULT1A1* gene and its relationship to other common *SULT1A1* variants. Furthermore, we demonstrate that the level of SULT1A1 enzyme activity is correlated most closely with copy number variation.

RESULTS

During a study in which pyrosequencing was used to genotype *SULT1A1*, an unusual peak distribution was observed for DNA samples heterozygous for the *SULT1A1*2* (Arg213His) polymorphism. After demonstrating that there were no polymorphisms under the primers, and that *SULT1A2* and *SULT1A3/4* were not co-amplified (data not shown), we considered the possibility that copy number differences for the *SULT1A1* gene might explain these observations. To test that hypothesis, a fluorescence-based semi-quantitative PCR assay was developed to co-amplify a portion of *SULT1A1* and *SULT1A2* as well as a fragment from the Coagulation Factor V gene. The location of the primers for *SULT1A1* and *SULT1A2* (labeled S1 and S2, respectively) is illustrated in Figure 1. By comparing peak heights for *SULT1A1* to the controls (*SULT1A2* and the Coagulation Factor V), copy number differences were detected. Figure 2 illustrates the results of this copy number assay for five separate patients showing the peak heights for *Factor 5* (F5), *SULT1A2* and multiple copies (1–5) of *SULT1A1* (upper panel). The lower panel illustrates the

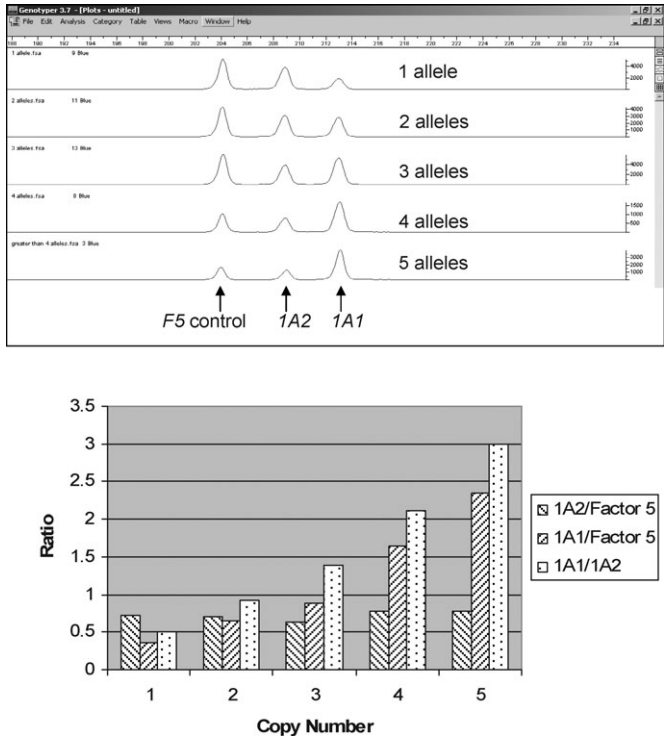

Figure 2. Results of the copy number assay for five separate individuals showing the peak heights for *Factor 5* (F5), *SULT1A2* and multiple copies (1–5) of *SULT1A1* (upper panel). The lower panel illustrates the ratios observed for 1A2/F5, 1A1/F5 and 1A1/1A2.

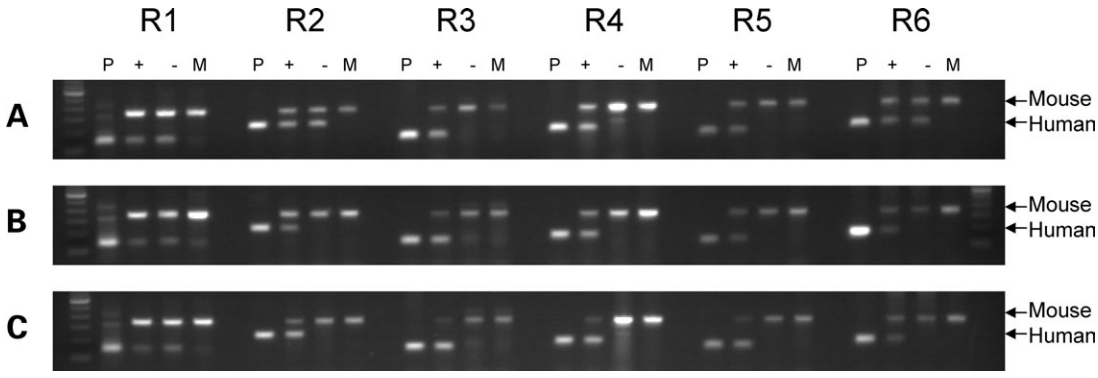

Figure 3. PCR analysis of three samples (A, B and C) subjected to haploid mapping. For each Parental sample (P), two haplotype cell lines were identified: chromosome 16 with intact *SULT1A1* (+) and chromosome 16 with deleted allele (−). In addition, mouse only DNA (M) was used as a control. Six PCR primer sets for regions R1 through R6 were utilized in a multiplex with mouse STS marker MEG1.

ratios observed for 1A2/F5, 1A1/F5 and 1A1/1A2. Since the 1A2 to F5 ratios were generally consistent among all samples tested, the 1A1/1A2 ratio was used to calculated copy number for the purposes of this study.

In order to define the extent of these genomic alterations, the region surrounding the *SULT1A1* gene was further analyzed in three cases found to have a deletion. Haploid cell lines were first created by fusing E2 mouse cells with lymphoblastoid cells obtained from these three subjects and then screened for clones containing only a single human chromosome 16. PCR analysis was then performed utilizing primers for six genomic regions (Fig. 1, labeled R1–R6). Results of the PCR deletion mapping for all three of these cases (DNA from the original lymphocytes or parental sample, the cell lines with intact SULT1A1 and the cell lines with deleted SULT1A1, along with mouse only DNA) is shown in Figure 3. On the basis of this analysis, at least two different deletions were identified. For sample A, the 3′ breakpoint of *SULT1A1* occurs in a 3.6 kb region between primers R2 and R3, while the 5′ breakpoint lies within a 3.1 kb region between R5 and R6. For both samples B and C, the 3′ deletion breakpoint occurs between R1 and R2 while the 5′ breakpoint extends beyond R6 (Figs 1 and 3).

Utilizing the quantitative assay described earlier, the frequency of *SULT1A1* copy number was then determined among 362 Caucasians (95 from the Coriell Cell Repository and 267 adult liver samples) and 99 African-American (Coriell Cell Repository) individuals. Results of the *SULT1A1* copy number assay for one of the groups tested (the 267 adult liver samples) is shown in Figure 4. A summary of results for all samples tested is provided in Table 1. Among all of the Caucasian subjects, 17 of 362 samples (4.7%) demonstrated a deletion within the *SULT1A1* gene, i.e. one copy, while 93 subjects (25.7%) had three or more copies. None of the African-American subjects carried a gene deletion while 62 of 99 subjects (62.6%) had three or more copies.

Three polymorphisms within the *SULT1A1* gene, G638A and the two 5′-FR polymorphisms, have been reported to play a role in determining the variability observed in *SULT1A1* enzymatic activity. In an effort to explore the functional significance of the two 5′-FR polymorphisms (−624

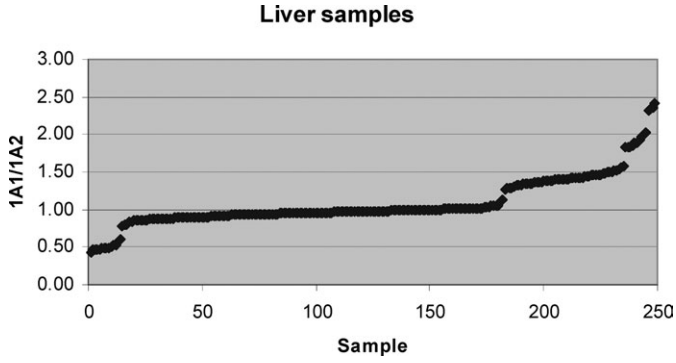

Figure 4. Results of the copy number assay showing the 1A1/1A2 ratio for each of the 267 adult liver samples tested. Those samples with a ratio of ∼0.5, 1.0, 1.5 and 2.0 were defined as having a *SULT1A1* copy number of 1, 2, 3 and 4, respectively.

and −396), luciferase reporter gene assays were performed with constructs containing the four possible haplotypes of the *SULT1A1* 5′-FR SNPs. After transfection into HepG2 and HEK293 cells, the CG, GG, CA and GA haplotypes showed a 8.3-fold, 7.5-fold, 3.8-fold and 0-fold ($n = 6–9$) increases in luciferase activity, respectively, when compared with pGL-3 Basic (Fig. 5). These results further support the idea that these polymorphic variants may contribute to differences in *SULT1A1* enzymatic activity.

Previous studies of SULT1A1 genotype–phenotype correlation, which have focused primarily on the G638A SNP and the two 5′-FR polymorphisms, have been performed in the absence of knowledge of the *SULT1A1* gene duplication. Thus, to further determine whether variations in the *SULT1A1* copy number and the three SNPs are associated with alteration in function, both copy number and phenotype for SULT1A1 was derived for 23 platelet samples and 267 liver samples; 18 (6%) of the 290 samples carried the deleted allele while 78 (27%) had more than two copies of the gene. Although copy number does not influence the accuracy of the genotyping calls for those samples with a copy number of 1 or 2 (∼75% of the samples), the accuracy is influenced for those cases with a copy number of 3 or more. Using

Table 1. Frequency of *SULT1A1* copy number differences in various study groups

Samples	1 allele	2 alleles	3 alleles	4 alleles	>4 alleles	>2 alleles
CA						
Coriell ($n = 95$), n(%)	2 (2.1)	73 (76.8)	17 (17.9)	3 (3.2)	0	20 (21.1)
Liver ($n = 267$), n(%)	15 (5.6)	179 (67.0)	59 (22.1)	10 (3.7)	4 (1.5)	73 (27.3)
Total ($n = 362$), n(%)	17 (4.7)	252 (69.6)	76 (21.0)	13 (3.6)	4 (1.1)	93 (25.7)
AA						
Coriell ($n = 99$), n(%)	0	37 (37.4)	37 (37.4)	21 (21.2)	4 (4.0)	62 (62.6)

CA, Caucasian-American; AA, African-American.

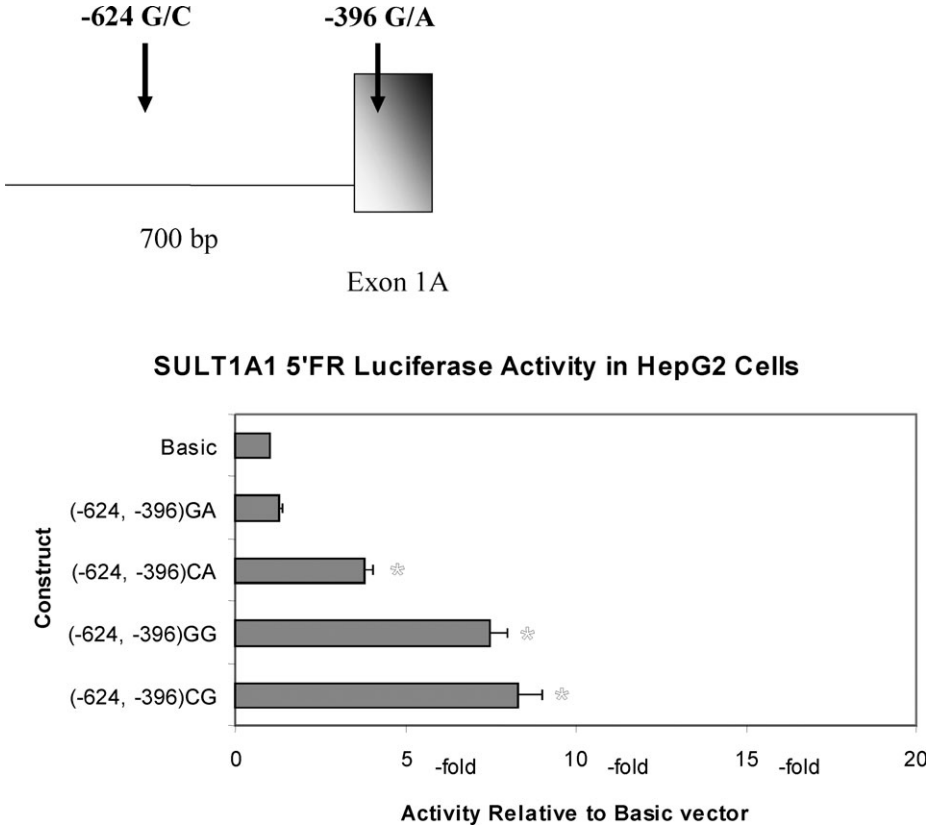

Figure 5. Reporter gene studies. Luciferase activity for 700 bp constructs for SNPs at both (−624) and (−396) transfected into HepG2 cells. Values represent mean ± SEM ($n = 6–9$) and * represents a significant difference in activity when compared with the basic vector, $P < 0.0001$.

the gene copy data in conjunction with a semi-quantitative SNP assay, the number of alleles for the G638A SNP and the two promoter SNPs were further defined in the liver samples. The effect of genotype versus the effect of copy number on level of activity was then determined as described in the methods section. Overall, there did not appear to be a significant effect of a particular allele above the effect of the total copy number on the level of activity for each of the three SNPs tested: 5′-FR (−624), rs3760091; 5′-FR (−396), rs750155 and G638A, rs9282861 ($P = 0.61$, 0.77 and 0.67, respectively) with respect to SULT1A1 activity. The total copy number, however, did have a very strong effect ($P < 0.0001$) with respect to enzyme activity. For each increase in copy number, there is an estimated 0.15 (95% CI: 0.11–0.19) associated increase in enzymatic activity ($\times 10^{-1}$ Units/mg) (Fig. 6A). A more pronounced effect was seen when enzyme activity in platelets was compared with *SULT1A1* copy number (Fig. 6B). The reason for this more pronounced effect in enzyme activity derived from platelets compared with liver, however, is not entirely clear.

DISCUSSION

It has been known for decades that the SULT1A1 level of activity is influenced by common genetic polymorphisms (10,11). However, the common coding region polymorphism, *SULT1A1*2* (G638A) is associated with only a portion of the variation observed in SULT1A1 activity (12,27). Recent

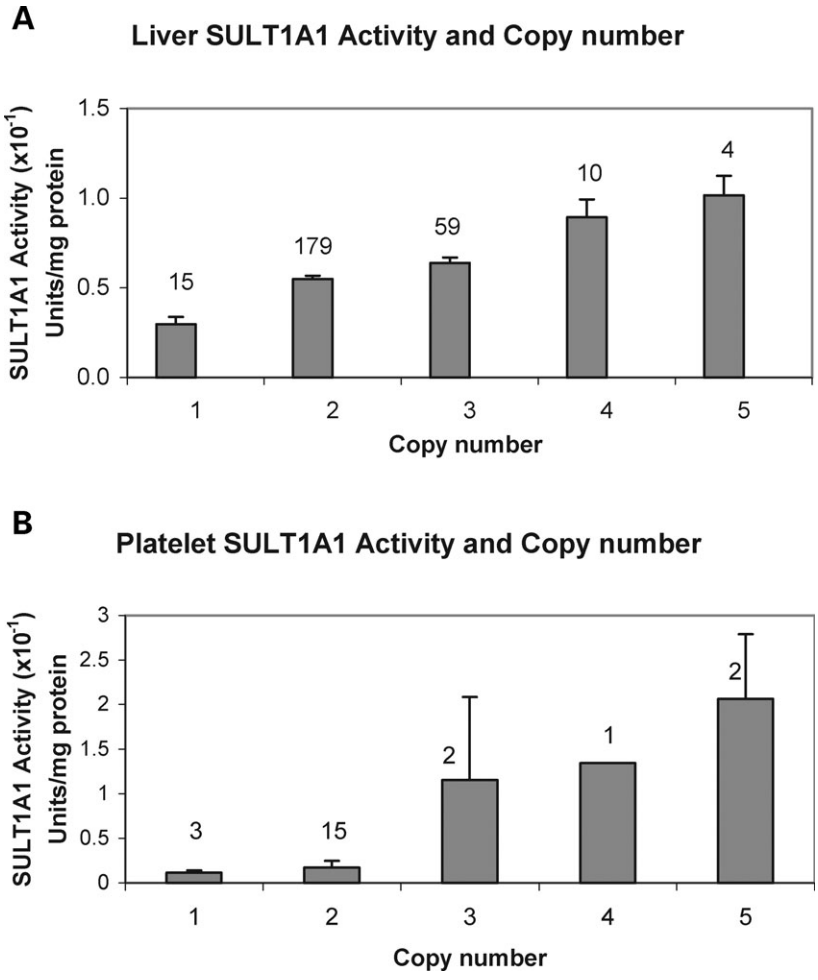

Figure 6. Plot of the enzyme activity level of SULT1A1 derived from different tissue as a function of copy number differences. (**A**) Enzymatic activity derived from 267 liver samples. (**B**) Activity derived from 23 selected platelet samples. The numbers above the bar represent the number of samples tested in each group.

studies have reported that additional polymorphisms in the *SULT1A1* 5′-FR and exon 1A, which are in linkage disequilibrium with *SULT1A1*2*, might also influence SULT1A1 phenotype (15,16). Reporter gene assays presented in this study support these *in vitro* findings. However, these common SNPs still do not completely explain variation in SULT1A1 phenotype *in vivo*. In the present study, we report that *SULT1A1* gene copy number differences occur and that copy number differences appear to explain the majority of the variation in SULT1A1 *in vitro* activity in liver and platelets when all of the presently known sources of genetic variability are taken into account. This finding most likely reflects the relative contribution of each of the polymorphic variants to the overall phenotype. That is, even though the reporter gene assays support the idea that the promoter polymorphic variants may contribute to differences in SULT1A1 enzymatic activity, their contribution to the enzymatic activity relative to the effect of copy number appears to be small.

Overall, these results may help to explain the inconsistencies found among those case–control studies testing for possible associations of *SULT1A1*2* polymorphism with risk for cancer (17–21). There are at least two potential sources of error in these and similar previous studies: (i) gene copy number has not been taken into account and (ii) incorrect genotypes due to copy number differences. Because of the presence of copy number differences, therefore, all previously reported genotypes results should be called into question.

The presence of multiple copies of *SULT1A1* appears to be common, with frequencies that vary with ethnicity. Twenty-six percent of Caucasians and 63% of the African-American subjects studied had more than two alleles. In addition, 5% of the Caucasian subjects had a single copy of the *SULT1A1* gene, while none of the African-American subjects tested had this genotype (Table 1). The fact that increased copy numbers for the *SULT1A1* were observed more often in African-American subjects may explain in part, the observation by Anderson *et al*. (28), that basal platelet SULT1A1 enzyme activity in African-American subjects is significantly higher than that in Caucasian subjects.

Data from the hybrid cell mapping demonstrate that the breakpoints flank the coding region and that the entire *SULT1A1* gene is deleted. Although not tested directly, we presume that gene duplications will have similar breakpoints. Because the 5′ and 3′ ends of *SULT1A1* share such high homology, the most likely mechanism for these deletion and duplication events is the presence of homologous recombination.

Other *SULT* genes localized to chromosome 16 are *SULT1A3* and *SULT1A4*, both ~800 Kb downstream of *SULT1A1*. Given the distance, we do not expect that these genes are involved in the deletion/duplication events detected in this study for *SULT1A1*. However, the extent of the chromosomal abnormality beyond *SULT1A1* is, at this time, completely unknown. Additionally, we cannot rule out the involvement of other genes that map more closely to *SULT1A1*.

In summary, we have documented the presence of a common gene duplication/deletion event in *SULT1A1*. Individuals carrying additional copies of *SULT1A1* represent 'rapid sulfators,' while those with fewer copies are 'slow sulfators'—as is the case for cytochrome p450 2D6, another drug metabolizing enzyme (25,29–31). The pharmacogenetic implications of differences in *SULT1A1* gene dosage might help explain individual differences in drug toxicity and/or efficacy in the clinical setting. Additionally, it should now be possible to consider *SULT1A1* gene dosage as an independent variable when studying the possible association of this gene with complex phenotypes and/or disease susceptibility. Finally, copy number differences must be taken into account when genotyping common SNPs within the *SULT1A1* gene.

METHODS

DNA

DNA for these studies was obtained from a variety of sources. One hundred DNA samples each from Caucasian-American and African-American subjects (100-item sample sets, HD100CAU and HD100AA), were obtained from the Coriell Cell Repository (Camden, NJ). DNA was also isolated from 23 subjects selected for extreme platelet SULT1A1 phenotypes (12). Finally, DNA was obtained from adult human hepatic surgical biopsy samples obtained from 268 Caucasian women having clinically indicated surgery at the Mayo Clinic, predominantly for the diagnosis and/or treatment of metastatic carcinoma. These samples were anonymized and only information with regard to diagnosis, gender, race and age was provided. Normal hepatic tissue from the site of tumor was used to perform the studies described subsequently. The Mayo Clinic Institutional Review Board reviewed and approved the use of these specimens for experimental purposes.

Enzyme activity

Platelet homogenates from 23 blood samples were previously phenotyped for SULT1A1 activity as described elsewhere (12). SULT1A1 enzyme activity in the liver samples were measured by modifications of the method of Foldes and Meek (32) as described by Campbell *et al*. (33).

SNP genotyping

Genotypes for the *SULT1A1*2* polymorphism (Arg213His, rs9282861) were obtained by pyrosequencing using a PSQ 96 instrument (Biotage) according to the manufacture's guidelines. Genotypes for −396 (rs750155) and −624 (rs3760091) were acquired using fluorescent-based allele-specific PCR with the analysis performed on an ABI3100 DNA analyzer (Applied Biosystems). The PCR conditions and primer sequences for all three of these polymorphisms are available upon request.

Copy number assay: fluorescent-based quantitative PCR

A set of PCR primers were designed to co-amplify a 212 bp fragment within exons 2 and 3 of *SULT1A1* (NM_177534) and a 208 bp fragment within exons 3 and 4 of *SULT1A2* (NM_177528) (Fig. 1, labeled S1 and S2, respectively). In this case, *SULT1A2* is used as an internal control for copy number. In addition, a 202 bp fragment amplified from the Coagulation Factor V gene was included as an additional copy number control. PCR products were analyzed on an ABI3100 DNA analyzer (Applied Biosystems) and copy number was estimated by calculating the height ratio of the 212 bp amplicon of *SULT1A1* to the reference 208 bp amplicon of *SULT1A2*. The PCR conditions and primer sequences for this copy number assay are available upon request.

Haploid analysis

EBV transformed lymphocytes of three samples known to carry a deletion of *SULT1A1* were subjected to haploid conversion. Hybrid cell lines were generated by fusing lymphoblastoid cells from subjects with E2 mouse cells essentially as described by Yan *et al*. (34). Using a number of microsatellite markers, mouse/human hybrid cell lines were screened to isolate clones containing individual chromosome 16 alleles: one clone containing the intact *SULT1A1* gene and another clone with the deleted allele. The extent of the deletion in each case was then examined directly by testing DNA for the presence or absence of PCR amplification using six primer pairs (labeled R1—R6) spaced along the 5′ and 3′ region of *SULT1A1* (Fig. 1). As a PCR control, a mouse STS marker (MEG1) was co-amplified in a multiplex with the six primer pairs. The PCR products were then run on a 2% agarose gel and visualized directly after staining with ethidium bromide. The PCR primer sequences and conditions are available upon request.

Reporter gene constructs and promoter activity

SULT1A1 5′-FR sequences that contain the common SNPs located at positions (−624) and (−396) were used to create firefly luciferase reporter gene constructs in pGL3-Basic vector (15). Nucleotide locations were numbered relative to the 'A' in the *SULT1A1* translation initiation codon. These constructs were 700 bp and contained all possible haplotypes for the two SNPs (−624 C/G, −396 G/A). These constructs were sequenced in both directions to ensure that the correct sequences were present and were then used to transfect HepG2 and HEK293 cells. Specifically, 2 μg purified plasmid DNA were transfected into the cells together with 20 ng pRL-TK (Promega) DNA. The *Renilla* luciferase activity expressed by pRL-TK was used as a control for transfection efficiency. Cells were also transfected with pGL3-Basic that lacked an insert as a control. Results of these reporter gene studies were reported as the ratio of firefly luciferase to *Renilla* luciferase light units, and values

were expressed relative to the activity of the pGL3-Basic vector construct. All assays were performed in triplicate in both cell lines and were repeated three times, for a total of nine independent data points.

Statistical analysis

To test the effect of genotype versus the effect of copy number on the level of activity, linear regression models were used. The first model was for the effect of copy number, regardless of allele type, and used the total number of allele copies as the independent predictor of activity level. The second model considered the role of particular allele types, and used the counts of each particular allele type as the predictors of activity level (i.e. requiring two independent predictors). Comparing the second model versus the first by a likelihood ratio test provided an evaluation of the role of particular alleles, over the effect of copy number, on level of activity. A small *P*-value would suggest that specific alleles influence activity level over the effect of total copy number, whereas a large *P*-value would suggest specific alleles do not contribute to activity level, once the total copy number is accounted for. Comparing the first model versus a model without any predictors provided a test of the total copy number.

ACKNOWLEDGEMENTS

Supported in part by National Institute of Health grants R01 GM28157, R01 GM35720 and U01 GM61388 (A.A.A., J.L.B., J.Z., R.M.W.) and by Prostate SPORE CA91956 (S.J.H., J.M.C., S.K.M., D.J.S., S.N.T.). Funding to pay the Open Access publication charges for this article was provided by the Department of Laboratory Medicine and Pathology.

Conflict of Interest statement. None declared.

REFERENCES

1. Falany, C.N. (1997) Enzymology of human cytosolic sulfotransferases. *FASEB J.*, **11**, 206–216.
2. Coughtrie, M.W. (2002) Sulfation through the looking glass—recent advances in sulfotransferase research for the curious. *Pharmacogenomics J.*, **2**, 297–308.
3. Glatt, H. and Meinl, W. (2004) Pharmacogenetics of soluble sulfotransferases (SULTs). *Naunyn Schmiedebergs Arch. Pharmacol.*, **369**, 55–68.
4. Hildebrandt, M.A., Salavaggione, O.E., Martin, Y.N., Flynn, H.C., Jalal, S., Wieben, E.D. and Weinshilboum, R.M. (2004) Human SULT1A3 pharmacogenetics: gene duplication and functional genomic studies. *Biochem. Biophys. Res. Commun.*, **321**, 870–878.
5. Weinshilboum, R.M. and Adjei, A.A. (2005) Sulfate conjugation: pharmacogenetics and pharmacogenomics. In Pacifici, G.M. and Coughtrie, M.W.H. (eds), *Human Cytosolic Sulfotransferases*, CRC Press, Boca Raton, FL, pp.61–78.
6. Her, C., Raftogianis, R. and Weinshilboum, R.M. (1996) Human phenol sulfotransferase STP2 gene: molecular cloning, structural characterization, and chromosomal localization. *Genomics*, **33**, 409–420.
7. Raftogianis, R.B., Her, C. and Weinshilboum, R.M. (1996) Human phenol sulfotransferase pharmacogenetics: STP1 gene cloning and structural characterization. *Pharmacogenetics*, **6**, 473–487.
8. Martin, J., Han, C., Gordon, L.A., Terry, A., Prabhakar, S., She, X., Xie, G., Hellsten, U., Chan, Y.M., Altherr, M. *et al*. (2004) The sequence and analysis of duplication-rich human chromosome 16. *Nature*, **432**, 988–994.
9. Adjei, A.A. and Weinshilboum, R.M. (2002) Catecholestrogen sulfation: possible role in carcinogenesis. *Biochem. Biophys. Res. Commun.*, **292**, 402–408.
10. Van Loon, J. and Weinshilboum, R.M. (1984) Human platelet phenol sulfotransferase: familial variation in thermal stability of the TS form. *Biochem. Genet.*, **22**, 997–1014.
11. Price, R.A., Spielman, R.S., Lucena, A.L., Van Loon, J.A., Maidak, B.L. and Weinshilboum, R.M. (1989) Genetic polymorphism for human platelet thermostable phenol sulfotransferase (TS PST) activity. *Genetics*, **122**, 905–914.
12. Raftogianis, R.B., Wood, T.C., Otterness, D.M., Van Loon, J.A. and Weinshilboum, R.M. (1997) Phenol sulfotransferase pharmacogenetics in humans: association of common SULT1A1 alleles with TS PST phenotype. *Biochem. Biophys. Res. Commun.*, **239**, 298–304.
13. Raftogianis, R.B., Wood, T.C. and Weinshilboum, R.M. (1999) Human phenol sulfotransferases SULT1A2 and SULT1A1: genetic polymorphisms, allozyme properties, and human liver genotype-phenotype correlations. *Biochem. Pharmacol.*, **58**, 605–616.
14. Carlini, E.J., Raftogianis, R.B., Wood, T.C., Jin, F., Zheng, W., Rebbeck, T.R. and Weinshilboum, R.M. (2001) Sulfation pharmacogenetics: SULT1A1 and SULT1A2 allele frequencies in Caucasian, Chinese and African-American subjects. *Pharmacogenetics*, **11**, 57–68.
15. Prondzinski, J., Thomae, B., Wang, L., Eckloff, B., Wieben, E. and Weinshilboum, R. (2003) Sulfotransferase (SULT) 1A1 pharmacogenetics: functional 5′-flanking region (5′-FR) polymorphisms. *Clin. Pharmacol. Ther.*, **73**, 77.
16. Ning, B., Nowell, S., Sweeney, C., Ambrosone, C.B., Williams, S., Miao, X., Liang, G., Lin, D., Stone, A., Ratnasinghe, D.L. *et al*. (2005) Common genetic polymorphisms in the 5′-flanking region of the SULT1A1 gene: haplotypes and their association with platelet enzymatic activity. *Pharmacogenet. Genomics*, **15**, 465–473.
17. Langsenlehner, U., Krippl, P., Renner, W., Yazdani-Biuki, B., Eder, T., Wolf, G., Wascher, T.C., Paulweber, B., Weitzer, W. and Samonigg, H. (2004) Genetic variants of the sulfotransferase 1A1 and breast cancer risk. *Breast Cancer Res. Treat.*, **87**, 19–22.
18. Saintot, M., Malaveille, C., Hautefeuille, A. and Gerber, M. (2003) Interactions between genetic polymorphism of cytochrome P450-1B1, sulfotransferase 1A1, catechol-o-methyltransferase and tobacco exposure in breast cancer risk. *Int. J. Cancer*, **107**, 652–657.
19. Han, D.F., Zhou, X., Hu, M.B., Wang, C.H., Xie, W., Tan, X.D., Zheng, F. and Liu, F. (2004) Sulfotransferase 1A1 (SULT1A1) polymorphism and breast cancer risk in Chinese women. *Toxicol. Lett.*, **150**, 167–177.
20. Zheng, L., Wang, Y., Schabath, M.B., Grossman, H.B. and Wu, X. (2003) Sulfotransferase 1A1 (SULT1A1) polymorphism and bladder cancer risk: a case–control study. *Cancer Lett.*, **202**, 61–69.
21. Nowell, S., Ratnasinghe, D.L., Ambrosone, C.B., Williams, S., Teague-Ross, T., Trimble, L., Runnels, G., Carrol, A., Green, B., Stone, A. *et al*. (2004) Association of SULT1A1 phenotype and genotype with prostate cancer risk in African-Americans and Caucasians. *Cancer Epidemiol. Biomarkers Prev.*, **13**, 270–276.
22. Conrad, D.F., Andrews, T.D., Carter, N.P., Hurles, M.E. and Pritchard, J.K. (2006) A high-resolution survey of deletion polymorphism in the human genome. *Nat. Genet.*, **38**, 75–81.
23. Hinds, D.A., Kloek, A.P., Jen, M., Chen, X. and Frazer, K.A. (2006) Common deletions and SNPs are in linkage disequilibrium in the human genome. *Nat. Genet.*, **38**, 82–85.
24. McCarroll, S.A., Hadnott, T.N., Perry, G.H., Sabeti, P.C., Zody, M.C., Barrett, J.C., Dallaire, S., Gabriell, S.B., Lee, C., Daly, M.J. *et al*. (2006) Common deletion polymorphisms in the human genome. *Nat. Genet.*, **38**, 86–92.
25. Johansson, I., Lundqvist, E., Bertilsson, L., Dahl, M.L., Sjoqvist, F. and Ingelman-Sundberg, M. (1993) Inherited amplification of an active gene in the cytochrome P450 CYP2D locus as a cause of ultrarapid metabolism of debrisoquine. *Proc. Natl Acad. Sci. USA*, **90**, 11825–11829.
26. Dalen, P., Dahl, M.L., Bernal Ruiz, M.L., Nordin, J. and Bertilsson, L. (1998) 10-Hydroxylation of nortriptyline in white persons with 0, 1, 2, 3 and 13 functional CYP2D6 genes. *Clin. Pharmacol. Ther.*, **63**, 444–452.
27. Nowell, S., Ambrosone, C.B., Ozawa, S., MacLeod, S.L., Mrackova, G., Williams, S., Plaxco, J., Kadlubar, F.F. and Lang, N.P. (2000) Relationship of phenol sulfotransferase activity (SULT1A1) genotype to sulfotransferase phenotype in platelet cytosol. *Pharmacogenetics*, **10**, 789–797.

28. Anderson, R.J., Jackson, B.L. and Liebentritt, D.K. (1988) Human platelet thermostable phenol sulfotransferase from blacks and whites: biochemical properties and variations in thermal stability. *J. Lab. Clin. Med.*, **112**, 773–783.
29. Gough, A.C., Miles, J.S., Spurr, N.K., Moss, J.E., Gaedigk, A., Eichelbaum, M. and Wolf, C.R. (1990) Identification of the primary gene defect at the cytochrome P450 CYP2D locus. *Nature*, **347**, 773–776.
30. Bertilsson, L., Dahl, M.L., Sjoqvist, F., Aberg-Wistedt, A., Humble, M., Johansson, I., Lundqvist, E. and Ingelman-Sundberg, M. (1993) Molecular basis for rational megaprescribing in ultrarapid hydroxylators of debrisoquine. *Lancet*, **341**, 63.
31. Gaedigk, A., Blum, M., Gaedigk, R., Eichelbaum, M. and Meyer, U.A. (1991) Deletion of the entire cytochrome P450 CYP2D6 gene as a cause of impaired drug metabolism in poor metabolizers of the debrisoquine/sparteine polymorphism. *Am. J. Hum. Genet.*, **48**, 943–950.
32. Foldes, A. and Meek, J.L. (1973) Rat brain phenolsulfotransferase: partial purification and some properties. *Biochim. Biophys. Acta*, **327**, 365–374.
33. Campbell, N.R., Van Loon, J.A. and Weinshilboum, R.M. (1987) Human liver phenol sulfotransferase: assay conditions, biochemical properties and partial purification of isozymes of the thermostable form. *Biochem. Pharmacol.*, **36**, 1435–1446.
34. Yan, H., Papadopoulos, N., Marra, G., Perrera, C., Jiricny, J., Boland, C.R., Lynch, H.T., Chadwick, R.B., de la Chapelle, A., Berg, K. *et al*. (2000) Conversion of diploidy to haploidy. *Nature*, **403**, 723–724.