Arch Toxicol (2017) 91:3329–3339
DOI 10.1007/s00204-017-1955-4

TOXICOGENOMICS

Methyleugenol DNA adducts in human liver are associated with *SULT1A1* copy number variations and expression levels

Roman Tremmel[1,2] · Kristin Herrmann[3,4] · Wolfram Engst[3] · Walter Meinl[3] · Kathrin Klein[1,2] · Hansruedi Glatt[3,4] · Ulrich M. Zanger[1,2]

Received: 13 January 2017 / Accepted: 9 March 2017 / Published online: 22 March 2017

Abstract Methyleugenol is a rodent hepatocarcinogen occurring in many herbs and spices as well as essential oils used for flavoring. Following metabolic activation by cytochromes P450 (CYPs) and sulfotransferases (SULTs), methyleugenol can form DNA adducts. Previously, we showed that DNA adduct formation by methyleugenol in mouse liver is dependent on SULT1A1 expression and that methyleugenol DNA adducts are abundant in human liver specimens. In humans, SULT1A1 activity is affected by genetic polymorphisms, including single-nucleotide polymorphisms (SNPs) and copy number variations (CNVs). Here we investigated the relationship between individual methyleugenol DNA adduct levels and SULT1A1 in human liver samples. Using isotope-dilution ultraperformance liquid chromatography coupled with tandem mass spectrometry, we quantified methyleugenol DNA adducts in 121 human surgical liver samples. Frequent CNVs, including deletions ($f=3.3\%$) and duplications ($f=36.4\%$) of *SULT1A1*, were identified using qPCR and TaqMan assays in the donors' genomic DNA. SULT1A1 mRNA and protein levels were quantified using microarray data and Western blot analysis, respectively. Methyleugenol DNA adducts were detected in all 121 liver samples studied. Their levels varied 122-fold between individuals and were significantly correlated to both mRNA and protein levels of SULT1A1 ($r_s=0.43$, and $r_s=0.44$, respectively). Univariate and multivariate statistical analysis identified significant associations of *SULT1A1* CNVs with mRNA ($p=1.7\times10^{-06}$) and protein ($p=4.4\times10^{-10}$) levels as well as methyleugenol DNA adduct levels ($p=0.003$). These data establish the importance of *SULT1A1* genotype for hepatic methyleugenol DNA adducts in humans, and they confirm a strong impact of *SULT1A1* CNVs on SULT1A1 hepatic phenotype.

Keyword Copy number variation · DNA adducts · Human liver · Methyleugenol · Pharmacogenetics/genomics · *SULT1A1* expression

R. Tremmel and K. Herrmann contributed equally to this study.

Electronic supplementary material The online version of this article (doi:10.1007/s00204-017-1955-4) contains supplementary material, which is available to authorized users.

✉ Ulrich M. Zanger
uli.zanger@ikp-stuttgart.de

1 Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology, Auerbachstr. 112, 70376 Stuttgart, Germany

2 University of Tuebingen, Tuebingen, Germany

3 German Institute of Human Nutrition (DIfE) Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany

4 Federal Institute for Risk Assessment (BfR), Berlin, Germany

Introduction

The alkylbenzene derivative methyleugenol (4-allyl-1,2-dimethoxybenzene) is a natural compound that has been detected in more than 450 plant species, many of them being used as foods, sources of essential oils or medicines (Tan and Nishida 2012). Particularly high levels of methyleugenol were observed in many herbs and spices, such as basil, pimento, laurel, anise, and nutmeg (De Vincenzi et al. 2000). Methyleugenol is also present in essential oils used for flavoring foods and cosmetics. In rats and mice, methyleugenol induces hepatocellular carcinomas, but also tumors in various other tissues, such as the bile duct, kidney, and

mammary gland (Miller et al. 1983; National Toxicology Program 2000). As a consequence, the usage of pure methyleugenol as an additive has been banned in some countries (e.g., European Union), but not in others (e.g., USA).

It has been known for quite some time that following metabolic activation, methyleugenol and other alkenylbenzenes can bind covalently to DNA and the resultant DNA adducts have been identified (Boberg et al. 1983; Phillips et al. 1984; Herrmann et al. 2012). The carcinogenicity of different alkenylbenzenes in the liver of rats and mice was shown to be directly associated with the abundance of DNA adducts (Miller et al. 1983; Phillips et al. 1984; Williams et al. 2013). Methyleugenol is metabolically activated in a two-step process. First, it is converted by various cytochrome P450 enzymes (CYPs), including human CYP1A2, CYP2C9, CYP2C19, and CYP2D6, to hydroxylated forms (in particular 1′-hydroxymethyleugenol) (Jeurissen et al. 2006; Al-Subeihi et al. 2015), which then acts as substrate for sulfotransferases (SULTs), especially members of the SULT1A family (Herrmann et al. 2012). The sulfated product 1′-sulfooxymethyleugenol is electrophilically reactive, and can form covalent bonds to the exocyclic amino groups of guanine and adenine bases in DNA, resulting in the formation of N^2-(*trans*-methylisoeugenol-3′-yl)-2′-deoxyguanosine (N^2-MIE-dG) and N^6-(*trans*-methylisoeugenol-3′-yl)-2′-deoxyadenosine (N^6-MIE-dA) adducts (Phillips et al. 1984; Herrmann et al. 2012). We previously showed that the SULT1A1 enzyme is the most efficient activator of 1′-hydroxymethyleugenol among the various human SULT forms expressed in *Salmonella typhimurium* (Herrmann et al. 2012). Likewise, hepatic DNA adduct formation by methyleugenol and 1′-hydroxymethyleugenol was decreased by 97 and 99%, respectively, in Sult1a1 knockout mice and strongly enhanced in mice transgenic for the human *SULT1A1-1A2* cluster (Herrmann et al. 2014). Moreover, DNA adduct formation in mice expressing endogenous Sult1a1 as well as human SULT1A1-1A2 was nearly additive compared to animals expressing only the endogenous enzyme or only the transgene, suggesting that SULT1A activity was limiting for the adduct formation. The *SULT1A1* status also had a strong impact on the methyleugenol-induced DNA adduct formation in several extrahepatic tissues (Herrmann et al. 2016).

Recently, we found that methyleugenol DNA adducts are abundant in human tissue samples. Using isotope-dilution ultraperformance liquid chromatography coupled with tandem mass spectrometry (UPLC-MS/MS), we detected these adducts in 29 out of 30 human liver samples (Herrmann et al. 2013) and in 10 out of 10 human lung samples (Monien et al. 2015). This method is particularly reliable with regard to the unambiguous structural identification and accurate quantification of DNA adducts (Tretyakova et al. 2012). Up to date, only a small number of DNA adducts from xenobiotics have been detected in human tissue samples by isotope-dilution UPLC-MS/MS or other robust methods. To the best of our knowledge, DNA adducts of no other xenobiotic have been observed in such a high percentage of the human tissue samples studied as with methyleugenol. Therefore, there is a need to clarify potential carcinogenic risks resulting from these adducts in humans. To this end, it may be useful to know factors determining this adduct formation. In the present study, we analyzed the role of some genetic and non-genetic host factors, in particular in relation to SULT1A1.

SULT1A1 is the most abundant SULT form in the human organism and ubiquitously expressed in human tissues, including the human liver (Teubner et al. 2007; Riches et al. 2009). Interindividual variability in SULT1A1 enzyme expression and activity is in part due to genetic factors, including functional genetic polymorphisms. For instance, a non-synonymous SNP *SULT1A1*2* (rs9282861, c.638G>A, Arg213His) reduces SULT1A1 enzyme activity in platelets (Raftogianis et al. 1997), and was associated to cancer risk in various studies (reviewed by Glatt and Meinl 2004; Sak and Everaus 2016). Moreover, structural variations, including copy number variations (CNVs), that alter functional gene dosage are common and have been associated with altered enzyme activity in platelets and in human liver (Hebbring et al. 2007; Yu et al. 2010; Gaedigk et al. 2012).

The aim of this study was to validate the role of the SULT1A1 enzyme in the formation of methyleugenol DNA adducts and to investigate the influence of genetic variance in the *SULT1A1* gene on SULT1A1 phenotypes and DNA adduct levels in human liver. To this end, we measured adduct levels in a well-documented cohort of 121 Caucasian liver samples using a previously established isotope-dilution UPLC-MS/MS procedure (Herrmann et al. 2013). We further quantified SULT1A1 expression levels and analyzed CNV and SNP genotypes of *SULT1A1*. Our data emphasize that the formation of methyleugenol DNA adducts in human liver depends on SULT1A1 activity and *SULT1A1* genotype. This finding should be useful for further studies on methyleugenol consumption in relation to cancer in humans.

Materials and methods

Liver samples

Liver tissue and corresponding blood samples for DNA extraction were previously collected from Caucasian patients undergoing liver surgery at the Department of General, Visceral, and Transplantation Surgery at the Charité

(Campus Virchow, University Medical Center Charité, Humboldt University Berlin, Germany). The use of human liver tissues was approved by the local ethical committees of the Charité, Humboldt University (Berlin, Germany), and the University Clinic Tuebingen, Germany. Written informed consent was obtained from all liver donors, and the study was conducted in accordance with the Declaration of Helsinki. All tissue samples had been examined by a pathologist, and only histologically normal liver tissue was collected and stored at −80 °C. In this study, we used 121 liver samples for which sufficient liver tissue material was available for the determination of methyleugenol DNA adducts and from which detailed clinical documentation was available (Gomes et al. 2009; Klein et al. 2010). All liver samples were from separate subjects, and there was no overlap with the liver samples analyzed previously for methyleugenol DNA adducts (Herrmann et al. 2013). Limited information was available on the dietary habits of the donors.

DNA samples and genotyping

Genomic DNA was isolated from EDTA blood samples using the QIAmp DNA Blood Mini Kit System (Qiagen, Hilden, Germany). The *SULT1A1* copy number was determined using commercially available TaqMan real-time quantitative PCR (qPCR) copy number assays (Applied Biosystems, Foster City, CA, USA) as specified in the manufacturer's instructions. In brief, 10 ng of genomic DNA was used in each reaction together with 2X TaqMan Universal Master Mix (Applied Biosystems) and TaqMan assay Hs04461762_cn together with VIC dye–labeled TaqMan RNase P assay (two gene copy reference) in a 7500 Real-Time PCR Instrument (Applied Biosystems). The absolute quantification was run in 12 µL reactions as quadruplicates and C_T values were recorded using manual C_T threshold and autobaseline. The *SULT1A1* copy number was quantified using CopyCaller™ software v2.0 (Applied Biosystems) and the comparative C_T ($\Delta\Delta$ C_T) method.

Genotyping for SNP rs9282861 was performed using restriction fragment-length (RFLP) analysis (Engelke et al. 2000). In brief, a genomic sequence [253 nucleotides (nt) from intron 6 and exon 7] containing the rs9282861 site was amplified by PCR using 70 ng of each DNA sample, the forward primer 5′-GCT CTG CAG GGT CTC TAG GA-3′ and the reverse primer 5′-ACG GTG GTG TAG TTG GTC ATA-3′. The resulting PCR product was treated with the restriction enzyme *BfoI* (Thermo Fisher Scientific, Braunschweig, Germany) and then electrophoresed on 3% agarose gels. The number of G and A alleles could be unambiguously determined from the height of the respective peaks following densitometric analysis of the gels using appropriate reference samples.

SULT1A1 mRNA expression

SULT1A1 mRNA levels were obtained using data previously collected by genome-wide expression array [IlluminaHuman WG6v2; SULT1A1 Probe_ID: ILMN_1656900; GEO: GSE32504 (Schröder et al. 2013)].

SULT1A1 protein expression

Relative SULT1A1 protein expression was analyzed using Western blotting of cytosolic fractions from human liver samples with adequate amounts of material ($n = 118$). Aliquots of 30 µg of total protein were separated by electrophoresis on 11% SDS–polyacrylamide gels and blotted onto nitrocellulose membranes. Sample loading and transfer efficiency were checked by Ponceau S staining. For the detection of SULT1A1 protein, anti-human SULT1A1 immunoserum raised in rabbit (1:10,000 dilution) and IRD800-labeled secondary anti-rabbit antibody (1:10,000 dilution, Licor, Lincoln, NE, USA) were used. The specificity of the antibody was confirmed using recombinantly expressed SULT1A forms (Supplementary Fig. S1) as well as other human SULT forms (Teubner et al. 2007). Immunoreactive bands were densitometrically quantified using an Odyssey infrared imaging system (Licor, Lincoln, NE, USA). For relative quantification, standard curves of pooled liver cytosols were co-analyzed on each blot (5, 10, 20, 40 µg protein) together with a reference sample. Values were normalized to the subject with the lowest SULT1A1 expression.

Detection of methyleugenol DNA adducts

The adducts were measured using isotope-dilution ultraperformance liquid chromatography (UPLC-MS/MS) as described previously (Herrmann et al. 2013). In brief, hepatic DNA was isolated from nuclear pellets using DNA extraction columns (Qiagen) and quantified using a NanodropND 1000 spectrophotometer (peQlab, Erlangen, Germany). Aliquots of 112.5 µg DNA were spiked with stable-isotope-labeled adduct standards ($[^{15}N_5]N^2$-MIE-dG and $[^{15}N_5]N^6$-MIE-dA) and digested to nucleosides using phosphodiesterase type II, micrococcal nuclease, and alkaline phosphatase. After DNA digestion and butanol extraction to reduce background signals, the samples were applied to the UPLC-MS/MS analysis. Since N^6-MIE-dA levels were determined to be approximately 50 times lower than N^2-MIE-dG levels, only the latter adducts were used for further analyses (Herrmann et al. 2012, 2013, 2014, 2016).

Statistical analyses

Associations and non-genetic factors Non-parametric Kruskal–Wallis rank sum tests, Wilcoxon rank sum tests, chi-squared tests, and Spearman's correlation tests were used to investigate the associations between SULT1A1 mRNA expression, protein expression or DNA adduct levels and the genetic factors (CNV, SNP) as well as demographic, clinical, and patient-related factors (Klein et al. 2010): age, sex [male vs female], nicotine [non-smokers vs moderate smokers (<1 packet of cigarettes/day) vs heavy smokers (>1 packet of cigarettes/day)], alcohol consumption [non-drinkers vs moderate drinkers (1–2 times/week) vs daily drinkers], presurgical drug exposure [no vs yes], cholestasis [no vs yes], serum bilirubin [normal (≤1.2) vs elevated (>1.2)], C-reactive protein [CRP; normal (≤8.2) vs elevated (>8.2)], serum gamma-glutamyltransferase [GGT; normal ($f \leq 36$; $m \leq 64$) vs elevated ($f > 36$; $m > 64$)], and diagnosis leading to liver resection [primary liver tumor vs metastasis].

ANOVA Multivariate median regression was applied to evaluate the association between *SULT1A1* CNVs and each phenotype corrected by the ten non-genetic factors. In brief, we used function anova.rq (with rank test-statistic and Wilcoxon score function) in R-package quantreg-5.19 (Koenker 2015) with the parameters rank test-statistic and Wilcoxon score to compare two median regression fits: (1) only the ten non-genetic factors as covariates versus (2) the ten genetic factors plus the *SULT1A1* CNV status (four groups; 1, 2, 3, ≥4).

Correlations and partial correlations Spearman's correlation tests and pairwise partial correlations function of the ppcor_1.1 package (Kim 2015) were applied as appropriate to investigate the correlation between DNA adduct measurements and each of the chosen gene expression levels. Expression profiles of 323 important genes involved in absorption, distribution, metabolism, and excretion (ADME) including genes of the core and extended ADME gene lists at pharmaadme.org and further genes responsible for xenobiotic metabolism were also extracted from expression array data (see above). In the partial correlation analysis, SULT1A1 mRNA expression levels were used as controlling variable. Reported p values were Benjamini Hochberg adjusted.

CNV × SNP interactions We adopted the regression model used by Hebbring et al. (2007) to assess whether the non-synonymous SNP *SULT1A1*2* had an influence above the CNV effect on both SULT1A1 phenotype levels and DNA adduct levels. For each phenotype and CNV (four groups; 1, 2, 3, ≥4) or SNP genotype (three groups; 0, 1, 2), we applied function anova.rq to compare different median regression fits. (1) A null model, (2) a model including either CNVs or SNPs alone, and (3) a third model including both CNV and SNP allele counts. We report pseudo F and p values, while a small p value (high F value) would suggest a more significant influence of the additional factor. All statistical tests were two-sided, and statistical significance was defined as $p < 0.05$. Analyses were performed using the statistic software R version 3.3.0 (R Core Team 2014).

Results

Descriptive analysis of DNA adduct levels

Methyleugenol DNA adducts (N^2-MIE-dG) were detectable in all 121 human liver samples using isotope-dilution UPLC-MS/MS. The adduct levels varied substantially between individual samples. The ratio between highest and lowest adduct measurements was 122, and the coefficient of variation as a measure of variability was 79%. We analyzed the contribution of ten available demographic and clinical host factors to the methyleugenol DNA adduct variability. As illustrated in Supplementary Figure S2, DNA adduct levels were slightly reduced in subjects with clinically elevated CRP levels, but this change was not significant ($p = 0.37$). Elevated serum bilirubin and GGT levels were significantly associated with lower DNA adduct levels (by 35 and 33%, respectively; both $p = 0.02$, pairwise Wilcoxon rank sum tests). Furthermore, there was a significant inverse relationship between the donor's age and the DNA adduct level, although the correlation coefficient was moderate ($r_s = -0.25$, $p = 0.006$).

ADME-wide correlation analyses

Taking advantage of genome-wide gene expression profiles that were available for the investigated liver samples from a previous microarray study (Schröder et al. 2013), we correlated the methyleugenol DNA adduct levels with the mRNA expression profiles of 323 pharmacogenes, including genes of the core and extended ADME gene lists at pharmaadme.org and further genes involved in xenobiotic metabolism. The analysis revealed a number of statistically significantly correlated genes (see Table 1 for the top correlated genes, and Supplementary Table S1 for the complete list). *SULT1A1* had the highest positive correlation coefficient, followed by several genes from the phase I and II enzyme, drug transporter as well as nuclear receptor families. To assess the contribution of these other genes in relation to *SULT1A1*, we used a partial correlation analysis with SULT1A1 mRNA expression as cofactor. This analysis revealed insignificant correlations between methyleugenol DNA adduct levels and most of the ADME genes except for *NR1H3*, coding for the liver X receptor LXR-A,

Table 1 Correlations between N^2-MIE-dG adduct levels and pharmacogene mRNA expression levels in 121 human liver samples

Gene	r_s	p^a	r (partial correlation)	P^a (partial correlation)
SULT1A1	0.426	**0.0003**	NA	NA
NR1H3	0.420	**0.0003**	0.344	**0.020**
CYP1A1	0.386	**0.001**	0.228	0.172
SULT1E1	0.384	**0.001**	0.264	0.121
NR3C1	0.383	**0.001**	0.239	0.150
SLCO1A2	0.379	**0.001**	0.282	0.083
CYP1A2	0.366	**0.002**	0.207	0.238
EPHX2	0.363	**0.002**	0.230	0.166
GSTZ1	0.355	**0.002**	0.192	0.306
PON1	0.350	**0.002**	0.221	0.189
RNF40	−0.290	**0.011**	−0.275	0.087
TRPV4	−0.295	**0.009**	−0.253	0.150
ABCB8	−0.297	**0.009**	−0.286	0.083
ABCC4	−0.301	**0.009**	−0.218	0.195
ABCC3	−0.305	**0.008**	−0.234	0.162
ABCB4	−0.322	**0.005**	−0.364	**0.015**
ALDH3B1	−0.331	**0.004**	−0.199	0.276
SLCO2A1	−0.351	**0.002**	−0.245	0.150
TYMS	−0.353	**0.002**	−0.298	0.083
GPX7	−0.356	**0.002**	−0.280	0.083

The top ten positively or negatively correlated genes are listed according to their Spearman correlation coefficients. In the partial correlation analysis SULT1A1 mRNA levels were used as controlling variable. The mRNA expression data were obtained from previous microarray analysis (Schröder et al. 2013)

[a]Benjamini Hochberg adjusted; significant p values (<0.05) are displayed in *bold* type

and *ABCB4*, coding for the transporter MDR3 (Table 1 and Supplementary Table S2).

SULT1A1 mRNA and protein expression

To further characterize SULT1A1 hepatic phenotypes, we determined SULT1A1 protein levels in the corresponding liver cytosols using an antiserum raised in rabbit against human SULT1A1 protein (Teubner et al. 2007). SULT1A1 mRNA and protein levels varied less strongly than the methyleugenol DNA adduct levels, as evident from their lower coefficients of variation of 30.3 and 27.6%, respectively. The highest and lowest SULT1A1 expression levels differed by a factor of 5.6 for mRNA and 4.5 for protein. SULT1A1 mRNA and protein levels were significantly correlated with each other (Fig. 1a; $r_s = 0.5$, $p = 1.0 \times 10^{-08}$). Although SULT1A1 mRNA levels were negatively correlated to the donor's age ($r_s = -0.19$, $p = 0.04$), the protein levels were not significantly influenced by the donor's age ($r_s = -0.11$, $p = 0.24$; Supplementary Fig. S3 and S4).

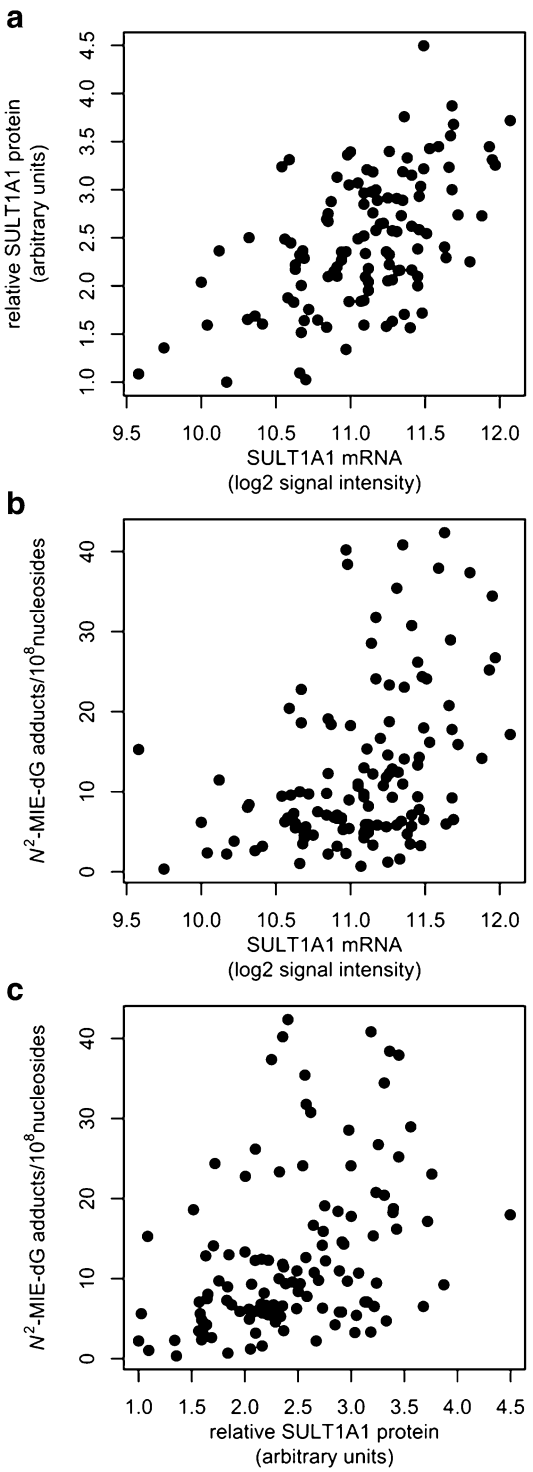

Fig. 1 Spearman correlation analysis between SULT1A1 phenotypes and methyleugenol DNA adduct (N^2-MIE-dG) levels in human liver donors. SULT1A1 mRNA expression levels were obtained from Illumina expression microarray data. Relative SULT1A1 protein expression was determined using Western blot analysis of cytosolic preparations. DNA adducts were measured in genomic DNA using isotope-dilution UPLC-MS/MS. The correlation coefficients and statistical significance were as follows: **a** $r_s = 0.5$, $p = 1.0 \times 10^{-08}$; **b** $r_s = 0.43$, $p = 1.1 \times 10^{-06}$; **c** $r_s = 0.44$, $p = 6.6 \times 10^{-07}$

Subjects diagnosed with metastases showed 1.15-fold higher SULT1A1 protein expression ($p=0.04$) compared to subjects with primary liver tumors (Supplementary Fig. S4).

As shown in Fig. 1b, c, the methyleugenol DNA adduct levels were correlated to both the SULT1A1 mRNA levels as well as the corresponding SULT1A1 protein levels. Both correlations were statistically highly significant ($r_s=0.43$, $p=1.1\times10^{-06}$; $r_s=0.44$, $p=6.6\times10^{-07}$).

Determination of *SULT1A1* copy number variation and frequencies in human liver

As illustrated in Fig. 2a, b, the copy number values observed could be assigned to five clusters representing different gene copy number groups. Four of 121 subjects with a calculated ratio below 1.4 were assigned copy number one. A group of 36 samples with a value >2.4 and <3.5 was assigned copy number three; a group of 7 donors with a value >3.5 and <4.5 was assigned copy number four, and finally one subject was assigned copy number five, as it exhibited a ratio >4.5. No donors were observed carrying a homozygous deletion. The calculated copy number values were all significantly different between the clusters ($p<0.001$; pairwise Wilcoxon rank sum tests). The observed CNV carrier frequencies were comparable to previous data, although we found slightly lower and higher frequencies compared to other samples with European ancestry for deletions and biallelic or multiallelic duplications, respectively (Table 2).

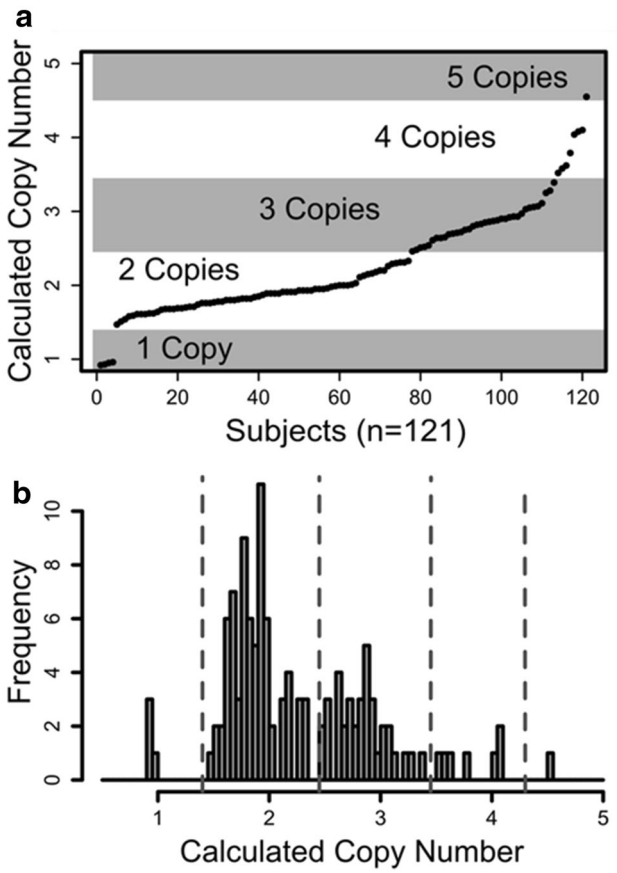

Fig. 2 *SULT1A1* gene copy number detection in 121 liver samples. The copy number was calculated using TaqMan C_T values and CopyCaller software v2.0 for *SULT1A1* intron 2 region. The distribution of copy number values within the cohort are shown as cluster diagram (**a**) and as histogram (**b**). *Gray and white bands* (**a**) and *vertical dashed lines* (**b**) define clusters of samples with different gene copy number

SULT1A1 genotype–phenotype correlation analysis

As illustrated in the boxplot in Fig. 3a, subjects with variant *SULT1A1* gene copy number showed proportionally altered mRNA expression levels ($p=4.4\times10^{-07}$, Kruskal–Wallis rank sum test). Whereas the decrease in the median mRNA level in subjects carrying one gene deletion was of borderline significance ($p=0.06$, pairwise Wilcoxon rank sum test), donors with three *SULT1A1* copies showed significantly increased median mRNA levels ($p=0.001$, pairwise Wilcoxon rank sum test) compared to donors with two gene copies. Subjects with four or more *SULT1A1* gene copies showed significantly higher median mRNA levels compared to donors with two ($p=1.0\times10^{-05}$) and three *SULT1A1* gene copies ($p=0.002$), respectively. The association between *SULT1A1* CNVs and mRNA expression remained significant in the multivariate analysis with correction for the non-genetic factors ($p=1.7\times10^{-06}$).

Analysis of relative SULT1A1 protein levels determined in liver cytosolic fractions revealed even stronger association with *SULT1A1* gene copy number both in a univariate test ($p=1.7\times10^{-11}$, Kruskal–Wallis rank sum test) as well as in the multivariate regression analysis ($p=4.4\times10^{-10}$). Subjects carrying a heterozygous *SULT1A1* gene deletion showed significantly lower median protein expression compared to samples with two gene copies (decrease by 31%; $p=0.04$, pairwise Wilcoxon rank sum test). The presence of three gene copies of *SULT1A1* significantly increased the median protein level by 1.4-fold ($p=4.2\times10^{-10}$, pairwise Wilcoxon rank sum test). The difference in protein

Table 2 Comparison of SULT1A1 CNV carrier frequencies in three studies

Study	Deletion (CN = 1) (%)	Duplication/s (CN > 2) (%)
Hebbring et al. (2007) ($n=362$)	4.7	25.7
Gjerde et al. (2008) ($n=151$)	5.3	29.9
This study ($n=121$)	3.3	36.4

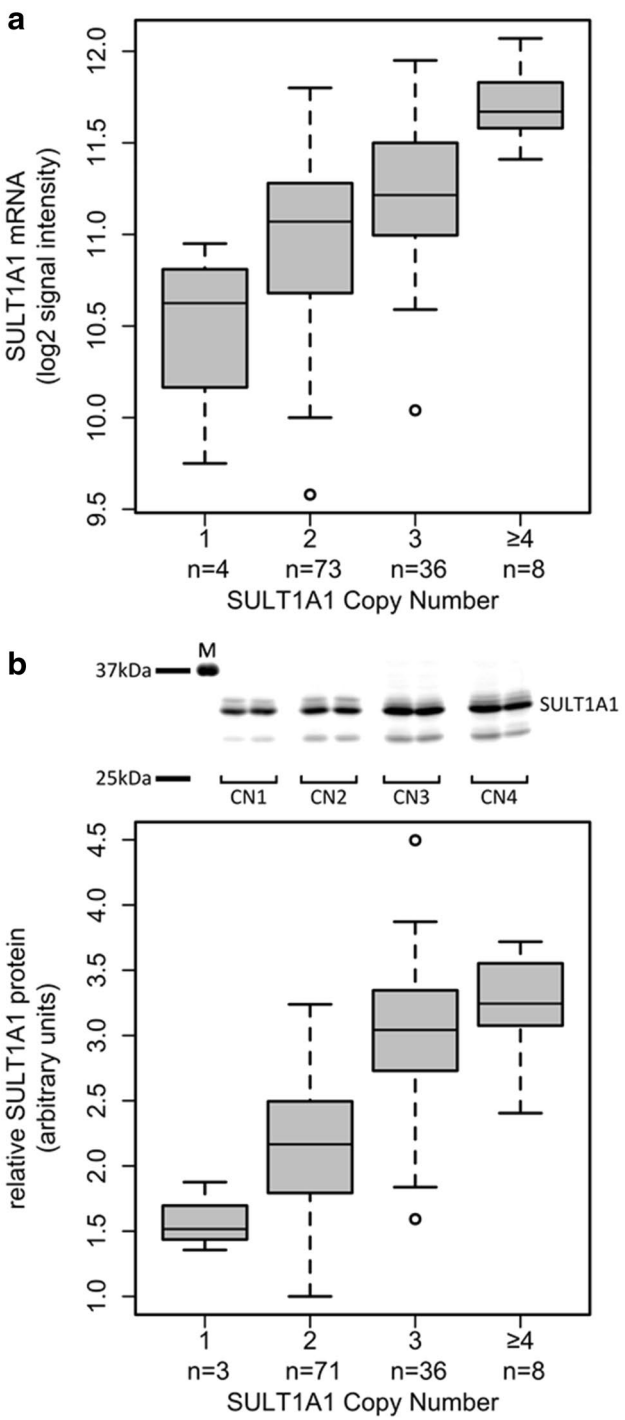

Fig. 3 Association between *SULT1A1* CNVs and SULT1A1 mRNA or relative protein expression. *SULT1A1* CNV status of 121 human liver samples was determined by qPCR and a TaqMan assay. **a** *Boxplots* displaying SULT1A1 mRNA levels for four distinct CNV groups. The corresponding mRNA expression levels were obtained from Illumina expression microarray data. **b** Representative and cropped Western blot showing SULT1A1 immunoreactive protein bands obtained with a polyclonal anti-SULT1A1 antibody in relation to CNV status (copy number one to four) for selected liver donors and a boxplot displaying SULT1A1 protein levels for four distinct CNV groups (one, two, three, and greater than or equal to four gene copies). Relative SULT1A1 protein expression was determined using Western blot analysis of 118 human liver cytosols. Three samples were not analyzed by Western blot analysis due to insufficient liver tissue

level between subjects carrying more than three *SULT1A1* gene copies compared to subjects with two copies was also significant ($p = 2.9 \times 10^{-05}$, pairwise Wilcoxon rank sum test) although less pronounced (Fig. 3b).

Influence of *SULT1A1* CNVs on DNA adduct levels

Subsequently we investigated the association between *SULT1A1* CNVs and methyleugenol DNA adduct (N^2-MIE-dG) levels in the 121 liver samples. As illustrated in the boxplot in Fig. 4, the level of DNA adducts increased consistently with the *SULT1A1* copy number ($p = 0.02$, Kruskal–Wallis rank sum test). Although subjects carrying one gene copy showed decreased DNA adduct levels by 25% compared to carriers of two *SULT1A1* gene copies, this difference was not statistically significant, probably due to the small number of samples. Significantly elevated adduct levels compared to subjects with normal CNV state of two copies were found for subjects carrying three *SULT1A1* gene copies (1.64-fold; $p = 0.03$, pairwise Wilcoxon rank sum test) and carriers of four or more gene copies (2.17-fold; $p = 0.01$, pairwise Wilcoxon rank sum test). The association of *SULT1A1* CNVs and DNA adduct levels was also significant in the multivariate analysis ($p = 0.003$).

Significance of the non-synonymous SNP *SULT1A1*2* (rs9282861, c.638G>A, Arg213His)

We genotyped the SNP rs9282861 in the liver samples to assess its functional significance for the SULT1A1 phenotypes measured and for the DNA adduct levels in relation to the *SULT1A1* CNVs. There appeared to be a consistent influence of the SNP genotype on mRNA levels, as these increased with the number of A alleles in each CNV group (not significant, Kruskal–Wallis rank sum test; Supplementary Fig. S5a). In contrast, no consistent influence was found for protein or methyleugenol DNA adduct levels. Donors carrying two gene copies with the heterozygous SNP showed decreased SULT1A1 protein expression levels (by 12%) compared to the homozygous non-carriers of the SNP (GG vs GA, $p = 0.01$, pairwise Wilcoxon rank sum test, Supplementary Fig. S5b). There were no further statistically significant associations of the SNP in relation to the CNV status on protein levels. The DNA adduct levels were not significantly associated to the A allele in any CNV group (Supplementary Fig. S5c). Furthermore, we used a median regression model to elucidate whether the SNP effect increases or compensates the CNV influence

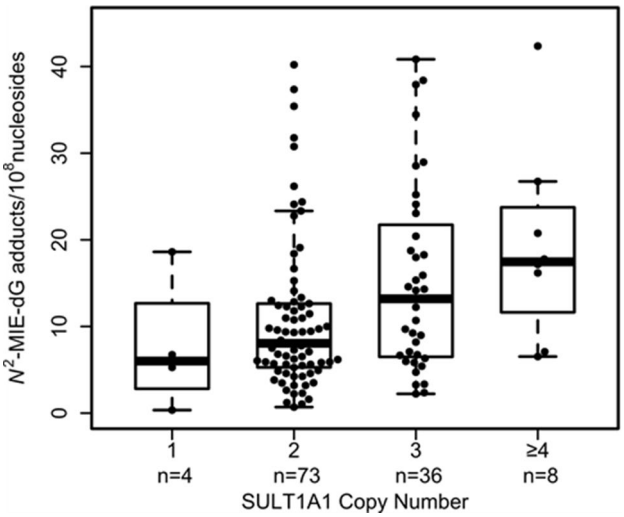

Fig. 4 Hepatic methyleugenol DNA adduct (N^2-MIE-dG) levels in relation to *SULT1A1* CNV status. Combination of box and scatter plots showing DNA adduct levels for four CNV groups of one, two, three, and greater than or equal to four gene copies. See legends of previous figures for methodological details

on SULT1A1 expression and methyleugenol DNA adduct measurements. This analysis revealed that the effect of the SNP on any phenotype was marginally significant compared to the CNV effect (Table S2). Only for the protein levels, there was a significant influence of the SNP alone ($p = 0.009$), which was however much weaker as compared to the *SULT1A1* CNVs (Table S2).

Discussion

In a previous study, we had established the dose–response curve for the formation of hepatic DNA adducts by methyleugenol in mouse models (Herrmann et al. 2014). In humanized mice (transgenic for *SULT1A1-1A2*, knockout of endogenous *Sult1a1*), DNA adducts were detected even at a single oral dose of 50 µg/kg body mass, which is four-fold below the estimated mean daily intake of methyleugenol by humans from foods, as estimated by a scientific committee of the European Union (2001). This finding led us to examine 30 human liver specimens for the presence of methyleugenol DNA adducts (Herrmann et al. 2013). The major adduct, N^2-MIE-dG, was unambiguously detected in 29 of the samples (at levels of 1.5–36.2 adducts/10^8 nucleosides), demonstrating that methyleugenol DNA adducts are abundant in human liver. Here, we analyzed a much larger set of 121 well-characterized liver samples. Because the liver cohort had originally been established for other purposes, no data on methyleugenol-related dietary habits of the donors were available. Nevertheless, it allowed us to investigate the population variability of hepatic methyleugenol DNA adducts and the influence of various genetic and non-genetic factors.

Increased age and elevated liver function parameters (serum GGT, serum bilirubin) correlated with lower hepatic methyleugenol DNA adduct levels. These relationships were independent of *SULT1A1* CNV status (data not shown). Possible explanations for the negative association with age could be reduced food intake or altered dietary preferences leading to reduced methyleugenol intake, or alterations in the expression/function of enzymes involved in the bioactivation or detoxification of methyleugenol. Furthermore, the enzyme activity of CYP1A2, which catalyzes the 1′-hydroxylation of methyleugenol (Jeurissen et al. 2006), is reduced in patients with elevated serum GGT or bilirubin (Klein et al. 2010). Since these serum parameters point to liver damage, liver cell proliferation may have been enhanced, leading to dilution of methyleugenol DNA adducts.

Because SULT1A1 was the known and primary candidate as methyleugenol adduct influencing factor, as based on work in genetically modified mouse models (Herrmann et al. 2014, 2016), we performed a comprehensive analysis of its hepatic expression variability as well as its genetic polymorphisms. These data considerably extend former studies of SULT1A1 in human liver, which have been more limited in extent and analysis (Hebbring et al. 2007; Riches et al. 2009). A polyclonal antibody raised against human SULT1A1 in sheep (Teubner et al. 2007) was used to study the SULT1A1 protein level in hepatic cytosolic fractions. This antibody readily reacts with all three human SULT1A forms (1A1, 1A2 and 1A3), but only shows minimal cross-reactivity with other human SULT forms (Teubner et al. 2007). We used recombinant SULT1A proteins, expressed in *S. typhimurium*, as standards to identify these forms in the human liver specimens. Confirming earlier reports (Heroux et al. 1989; Riches et al. 2009), SULT1A3 protein was not detected in the liver cytosols. Data on SULT1A2 protein in human liver are scarce in the literature. In part, this may be due to its great similarity to SULT1A1. The reference type forms (*1) of SULT1A1 and SULT1A2 have molecular masses of 34,195 and 34,308 Da, respectively, and only differ in 12 out of 295 amino acid residues, making electrophoretic separation or selective immunodetection difficult. Our conditions are optimized in this regard (Teubner et al. 2007). In the present study, we detected SULT1A2 protein, a modest activator of 1′-hydroxymethyleugenol (Herrmann et al. 2012), in all samples with adequate liver material (118 of 121 samples), but always at a markedly lower level (<1/4) than that of the SULT1A1 protein, similar to the findings in our previous studies (Meinl et al. 2006; Teubner et al. 2007).

The highest and lowest SULT1A1 protein levels differed by a factor of 4.5, which is markedly less than in the

study of Riches et al. (2009) with 11.7-fold variation in 28 liver samples. The SULT1A1 mRNA level correlated with the SULT1A1 protein level ($p = 1.0 \times 10^{-08}$). Both expression parameters correlated with the methyleugenol DNA adduct levels with overwhelming statistical significance ($r_s = 0.43$, $p = 1.1 \times 10^{-06}$; $r_s = 0.44$, $p = 6.6 \times 10^{-07}$, respectively), although their interindividual variability was much lower than that of the adducts (5.6- and 4.5-fold, compared to 122-fold). Among the available clinical factors tested, cholestasis, serum liver function parameters, age (reduced mRNA in elder subjects), and diagnosis leading to surgery (increased SULT1A1 protein levels in patients with metastases compared to patients with primary liver tumors) resulted in statistically significant but modest changes in SULT1A1 expression.

Using TaqMan qPCR analysis, we observed frequent *SULT1A1* CNVs in the 121 human liver samples. Compared to 36.4% of subjects carrying more than two gene copies, deletions were observed with lower frequencies (3.3%). The frequencies calculated in this study were consistent with previous studies, emphasizing their relatively common occurrence (Hebbring et al. 2007; Gaedigk et al. 2012). While earlier studies have shown the effect of *SULT1A1* CNVs on mRNA and enzyme activity (Hebbring et al. 2007; Yu et al. 2010), our data further expand the *SULT1A1* copy number–phenotype relationships by including quantitative protein expression data. Remarkably, we found the *SULT1A1* copy number polymorphism to be highly penetrant and to alter not only SULT1A1 expression, but also the abundance of methyleugenol DNA adducts in human liver, which clearly depended on the *SULT1A1* copy number status in this study. Moreover, median adduct levels increased almost linearly with *SULT1A1* gene copy number. Thus, carriers of four copies of *SULT1A1* per diploid cell showed almost three-fold higher methyleugenol DNA adduct levels compared to carriers with a single gene copy. The DNA adduct differences between carriers of more than two gene copies compared to carriers of two copies were significant by both univariate and multivariate analysis. The lack of significance for the difference between one and two copies was probably due to the smaller number of samples carrying a gene deletion. Our findings are consistent with previous work using a transgenic mouse model mimicking multiple *SULT1A1* gene copies (Hermann et al. 2014). However, given the complexity of metabolic and DNA repair mechanisms involved, we think it is quite surprising that the *SULT1A1* copy number polymorphism affects methyleugenol DNA adduct formation to an extent detectable in a cohort even without correction for (presumably high) variation of exposure. The inclusion of allele *SULT1A1*2* in a combined analysis of CNV and SNP genotypes revealed that CNVs had a much stronger impact on mRNA and protein expression as well as methyleugenol DNA adduct levels than the allele *SULT1A1*2*.

In order to obtain further insight into influential genes for adduct formation in a systematic manner, we performed correlation analyses based on mRNA expression profiles of over 300 ADME genes (Table 1; Supplemental Table S1). This approach showed that *SULT1A1* was the gene with the highest correlation coefficient and statistical significance. However, these values were only slightly lower for a number of other genes, including genes already known to be involved in methyleugenol metabolism, such as other SULT forms (*SULT1A2, SULT1E1*) and cytochrome P450s *CYP1A1, CYP1A2, CYP2C19* (Jeurissen et al. 2006; Herrmann et al. 2012, 2013). However, it should be noted that none of the co-regulated enzyme genes stayed significant in our partial correlation analysis, which corrects for the extent of co-regulation with SULT1A1. The gene with the second highest correlation coefficient, which was significant by both correlation strategies, was identified as *NR1H3*, a nuclear receptor coding for the liver X receptor alpha (LXRα). It is involved in the regulation of many components of the xenobiotic-metabolizing system. To our knowledge, a role of LXRα in the transcriptional regulation of SUL1A1 has not been established. *ABCB4* was the only other gene retaining statistically significant correlation with methyleugenol adducts after correction for co-regulation with SULT1A1. It encodes MDR3, a transmembrane transporter for lipids. It is difficult to imagine a direct role of MDR3 in the distribution and metabolism of methyleugenol.

In conclusion, we confirmed the strong influence of *SULT1A1* CNVs on hepatic SULT1A1 phenotype, and we presented a significant association to methyleugenol DNA adduct formation. Further studies are required to investigate a cancer risk of methyleugenol in relation to the uptake of the natural compound and *SULT1A1* copy number status in humans.

Acknowledgements The study was supported by the Robert Bosch Foundation, Stuttgart, Germany, and the German Federal Institute for Risk Assessment, Berlin, Germany (Grant FK-3-1329.421.1 + 2). R.T. was partially supported by a European Commission Horizon 2020 UPGx grant (668353). Britta Klumpp and Igor Liebermann are gratefully acknowledged for excellent technical assistance. We also thank Dr. Stefan Winter for expert advice concerning statistical analysis.

Compliance with ethical standards

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Conflict of interest The authors declare that they have no conflict of interest.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Al-Subeihi AAA, Alhusainy W, Kiwamoto R, Spenkelink B, van Bladeren PJ, Rietjens IMCM, Punt A (2015) Evaluation of the interindividual human variation in bioactivation of methyleugenol using physiologically based kinetic modeling and Monte Carlo simulations. Toxicol Appl Pharmacol 283:117–126. doi:10.1016/j.taap.2014.12.009

Boberg EW, Miller EC, Miller JA, Poland A, Liem A (1983) Strong evidence from studies with brachymorphic mice and pentachlorophenol that 1′-sulfoöxysafrole is the major ultimate electrophilic and carcinogenic metabolite of 1′-hydroxysafrole in mouse liver. Cancer Res 43:5163–5173

De Vincenzi M, Silano M, Stacchini P, Scazzocchio B (2000) Constituents of aromatic plants: I. Methyleugenol. Fitoterapia 71:216–221. doi:10.1016/S0367-326X(99)00150-1

Engelke CE, Meinl W, Boeing H, Glatt H (2000) Association between functional genetic polymorphisms of human sulfotransferases 1A1 and 1A2. Pharmacogenetics 10:163–169

European Union (2001) Opinion of the Scientific Committee on Food on Methyleugenol (4-Allyl-1,2-dimethoxybenzene). https://ec.europa.eu/food/sites/food/files/safety/docs/sci-com_scf_out102_en.pdf. Accessed 10 Feb 2017

Gaedigk A, Twist GP, Leeder JS (2012) CYP2D6, SULT1A1 and UGT2B17 copy number variation: quantitative detection by multiplex PCR. Pharmacogenomics 13:91–111. doi:10.2217/pgs.11.135

Gjerde J, Hauglid M, Breilid H, Lundgren S, Varhaug JE, Kisanga ER, Mellgren G, Steen VM, Lien EA (2008) Effects of CYP2D6 and SULT1A1 genotypes including SULT1A1 gene copy number on tamoxifen metabolism. Ann Oncol 19:56–61. doi:10.1093/annonc/mdm434

Glatt H, Meinl W (2004) Pharmacogenetics of soluble sulfotransferases (SULTs). Naunyn Schmiedebergs Arch Pharmacol 369:55–68. doi:10.1007/s00210-003-0826-0

Gomes AM, Winter S, Klein K, Turpeinen M, Schaeffeler E, Schwab M, Zanger UM (2009) Pharmacogenomics of human liver cytochrome P450 oxidoreductase: multifactorial analysis and impact on microsomal drug oxidation. Pharmacogenomics 10:579–599. doi:10.2217/pgs.09.7

Hebbring SJ, Adjei AA, Baer JL, Jenkins GD, Zhang J, Cunningham JM, Schaid DJ, Weinshilboum RM, Thibodeau SN (2007) Human SULT1A1 gene: copy number differences and functional implications. Hum Mol Genet 16:463–470. doi:10.1093/hmg/ddl468

Heroux JA, Falany CN, Roth JA (1989) Immunological characterization of human phenol sulfotransferase. Mol Pharmacol 36:29–33

Herrmann K, Engst W, Appel KE, Monien BH, Glatt H (2012) Identification of human and murine sulfotransferases able to activate hydroxylated metabolites of methyleugenol to mutagens in *Salmonella typhimurium* and detection of associated DNA adducts using UPLC-MS/MS methods. Mutagenesis 27:453–462. doi:10.1093/mutage/ges004

Herrmann K, Schumacher F, Engst W, Appel KE, Klein K, Zanger UM, Glatt H (2013) Abundance of DNA adducts of methyleugenol, a rodent hepatocarcinogen, in human liver samples. Carcinogenesis. doi:10.1093/carcin/bgt013

Herrmann K, Engst W, Meinl W, Florian S, Cartus AT, Schrenk D, Appel KE, Nolden T, Himmelbauer H, Glatt H (2014) Formation of hepatic DNA adducts by methyleugenol in mouse models: drastic decrease by Sult1a1 knockout and strong increase by transgenic human SULT1A1/2. Carcinogenesis 35:935–941. doi:10.1093/carcin/bgt408

Herrmann K, Engst W, Florian S, Lampen A, Meinl W, Glatt HR (2016) The influence of the SULT1A status—wild-type, knockout or humanized—on the DNA adduct formation by methyleugenol in extrahepatic tissues of mice. Toxicol Res 5:808–815. doi:10.1039/C5TX00358J

Jeurissen SMF, Bogaards JJP, Boersma MG, ter Horst JPF, Awad HM, Fiamegos YC, van Beek TA, Alink GM, Sudhölter EJR, Cnubben NHP, Rietjens IMCM (2006) Human cytochrome p450 enzymes of importance for the bioactivation of methyleugenol to the proximate carcinogen 1′-hydroxymethyleugenol. Chem Res Toxicol 19:111–116. doi:10.1021/tx050267h

Kim S (2015) Ppcor: an r package for a fast calculation to semi-partial correlation coefficients. Commun Stat Appl Methods 22:665–674. doi:10.5351/CSAM.2015.22.6.665

Klein K, Winter S, Turpeinen M, Schwab M, Zanger UM (2010) Pathway-Targeted Pharmacogenomics of CYP1A2 in Human Liver. Front Pharmacol. doi:10.3389/fphar.2010.00129

Koenker R (2015) Quantreg: quantile regression. R package version 5.11. http://CRAN.R-project.org/package=quantreg

Meinl W, Pabel U, Osterloh-Quiroz M, Hengstler JG, Glatt H (2006) Human sulphotransferases are involved in the activation of aristolochic acids and are expressed in renal target tissue. Int J Cancer 118:1090–1097. doi:10.1002/ijc.21480

Miller EC, Swanson AB, Phillips DH, Fletcher TL, Liem A, Miller JA (1983) Structure-activity studies of the carcinogenicities in the mouse and rat of some naturally occurring and synthetic alkenylbenzene derivatives related to safrole and estragole. Cancer Res 43:1124–1134

Monien BH, Schumacher F, Herrmann K, Glatt H, Turesky RJ, Chesné C (2015) Simultaneous detection of multiple DNA adducts in human lung samples by isotope-dilution UPLC-MS/MS. Anal Chem 87:641–648. doi:10.1021/ac503803m

National Toxicology Program (2000) NTP Toxicology and Carcinogenesis Studies of Methyleugenol (CAS NO. 93-15-2) in F344/N Rats and B6C3F1 Mice (Gavage Studies). Natl Toxicol Program Tech Rep Ser 491:1–412

Phillips DH, Reddy MV, Randerath K (1984) 32P-Post-labelling analysis of DNA adducts formed in the livers of animals treated with safrole, estragole and other naturally-occurring alkenylbenzenes. II. Newborn male B6C3F1 mice. Carcinogenesis 5:1623–1628. doi:10.1093/carcin/5.12.1623

R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/

Raftogianis RB, Wood TC, Otterness DM, Van Loon JA, Weinshilboum RM (1997) Phenol sulfotransferase pharmacogenetics in humans: association of common SULT1A1 alleles with TS PST phenotype. Biochem Biophys Res Commun 239:298–304. doi:10.1006/bbrc.1997.7466

Riches Z, Stanley EL, Bloomer JC, Coughtrie MWH (2009) Quantitative evaluation of the expression and activity of five major Sulfotransferases (SULTs) in human tissues: the SULT "Pie". Drug Metab Dispos 37:2255–2261. doi:10.1124/dmd.109.028399

Sak K, Everaus H (2016) Sulfotransferase 1A1 as a biomarker for susceptibility to carcinogenesis: from molecular genetics to the role of dietary flavonoids. Curr Drug Metab 17:528–541. doi:10.2174/1389200217666160219113924

Schröder A, Klein K, Winter S, Schwab M, Bonin M, Zell A, Zanger UM (2013) Genomics of ADME gene expression: mapping expression quantitative trait loci relevant for absorption,

distribution, metabolism and excretion of drugs in human liver. Pharmacogenomics J 13:12–20. doi:10.1038/tpj.2011.44

Tan KH, Nishida R (2012) Methyl eugenol: its occurrence, distribution, and role in nature, especially in relation to insect behavior and pollination. J Insect Sci. doi:10.1673/031.012.5601

Teubner W, Meinl W, Florian S, Kretzschmar M, Glatt H (2007) Identification and localization of soluble sulfotransferases in the human gastrointestinal tract. Biochem J 404:207–215. doi:10.1042/BJ20061431

Tretyakova N, Goggin M, Sangaraju D, Janis G (2012) Quantitation of DNA adducts by stable isotope dilution mass spectrometry. Chem Res Toxicol 25:2007–2035. doi:10.1021/tx3002548

Williams GM, Iatropoulos MJ, Jeffrey AM, Duan J-D (2013) Methyleugenol hepatocellular cancer initiating effects in rat liver. Food Chem Toxicol 53:187–196. doi:10.1016/j.fct.2012.11.050

Yu X, Dhakal IB, Beggs M, Edavana VK, Williams S, Zhang X, Mercer K, Ning B, Lang NP, Kadlubar FF, Kadlubar S (2010) Functional genetic variants in the 3′-untranslated region of sulfotransferase isoform 1A1 (SULT1A1) and their effect on enzymatic activity. Toxicol Sci Off J Soc Toxicol 118:391–403. doi:10.1093/toxsci/kfq296