GSTMI copy number and promoter haplotype as predictors for risk of recurrence and/or second primary tumor in patients with head and neck cancer

Xuemei Zhang\textsuperscript{1}
Maosheng Huang\textsuperscript{2}
Xifeng Wu\textsuperscript{2}
Susan Kadlubar\textsuperscript{1}
Jie Lin\textsuperscript{2}
Xinfeng Yu\textsuperscript{1}
Chunyang Fan\textsuperscript{3}
Baitang Ning\textsuperscript{4}
Fred F Kadlubar\textsuperscript{1,†}

\textsuperscript{1}University of Arkansas for Medical Sciences, Little Rock, Arkansas,
\textsuperscript{2}The University of Texas MD Anderson Cancer Center, Houston, Texas, \textsuperscript{3}VA hospital, Little Rock, Arkansas, \textsuperscript{4}National Center for Toxicological Research, US Food and Drug Administration, Jefferson, Arkansas, USA

\textsuperscript{†}Fred F Kadlubar passed away on December 4, 2010.

\textbf{Abstract:} The objective of this study was to determine copy number variant (CNV) and promoter genetic variants in glutathione S-transferase Mu class 1 (\textit{GSTM1}) and the risk of recurrence (REC)/second primary tumor (SPT) in patients with previously diagnosed early stage head and neck cancer. Among 441 subjects, 133 experienced REC and/or an SPT, while 308 had single primary disease. TaqMan real-time polymerase chain reaction was used to measure the exact copy number of \textit{GSTM1} and direct sequencing was used to determine genetic variants in the \textit{GSTM1} promoter region. Multivariate Cox regression analysis was performed to estimate hazard ratios (HRs) and 95% confidence intervals (95% CIs) associated with copy number and genetic variants. REC/SPT-free survival times were compared by constructing Kaplan–Meier curves and differences between curves were tested by logrank test. Results showed a significantly decreased REC/SPT (HR = 0.57; 95% CI = 0.35–0.95) and longer REC/SPT-free survival in subjects with at least two copies of \textit{GSTM1} compared with the \textit{GSTM1} homozygous deletion, but not in those with one copy of \textit{GSTM1}. The \textit{−498G, −426G, and −339T} alleles were significantly associated with REC/SPT, with HRs of 0.11 (0.02–0.85), 0.28 (0.11–0.74) and 2.02 (1.07–3.82), respectively. Kaplan–Meier survival analysis showed that the \textit{−498G, −426G, and −339C} alleles were also significantly associated with increased REC/SPT-free survival. Further haplotype analysis showed the haplotype P\textsuperscript{−498G−426G−339C} carriers had decreased REC/SPT with a HR of 0.09 (95% CI 0.01–0.71) and increased REC/SPT-free survival compared with those with haplotype P\textsuperscript{−498C−426A−339T}. The P\textsuperscript{−498C−426A−339T}-containing reporter construct had significantly increased luciferase expression. These results suggest that the \textit{GSTM1} CNV and promoter haplotype are better predictors of REC/SPTs of head and neck cancer than just measuring the presence/absence of \textit{GSTM1}.

\textbf{Keywords:} \textit{GSTM1}, copy number variant, REC, SPT, single nucleotide polymorphism

\section*{Introduction}

Head and neck cancer (HNC), which includes carcinomas of the oral cavity, pharynx, and larynx, is one of the most common human cancers worldwide. HNC accounts for 3–5\% of all cancers in the USA and it has been estimated that in 2012 more than 52,000 (40,250 oral and pharynx and 12,360 larynx) individuals will be diagnosed with HNC and 11,500 (7850 oral and pharynx and 3650 larynx) will die of the disease.\textsuperscript{1} Treatment of HNC is limited in early stage disease to either radiotherapy or surgery.\textsuperscript{2} However, post-stage treatment morbidity and mortality increases when patients experience recurrence of disease (REC) or develop a second primary tumor (SPT).\textsuperscript{3}
SPTs and recurrences develop in up to 20% of patients within 5 years of curative treatment. Smoking and alcohol consumption are strongly associated with poor outcomes in HNC. In addition, inherited factors, including genetic variants at specific genes, have also been demonstrated to modify this risk.

Glutathione S-transferase Mu (GSTM) class enzymes, members of the glutathione S-transferase (GST) superfamily of phase II drug-metabolizing enzymes, play an important role in protecting cells against xenobiotics by conjugating with glutathione to detoxify electrophilic compounds. The genes encoding the GSTM family are located on chromosome 1p13.3 in the order GSTM4, GSTM2, GSTM1, GSTM5, GSTM3. GSTM proteins have distinct tissue distribution: GSTM1 is a major Mu-class GST in the liver, the GSTM2 subunit is primarily associated with skeletal muscle, GSTM3 is enriched in testis, and GSTM5 is found in brain. It is well known that hepatic GSTM1 is highly polymorphic, and these genetic variations, which include copy number variants (CNVs) and common single nucleotide polymorphisms (SNPs), are likely to contribute to inter-individual differences in response to carcinogens and drugs. The homozygous deletion of the GSTM1 gene, which results in the absence of the GSTM1 enzyme, is present in 48%–57% of Caucasians, 23%–41% of African Americans, 32%–53% of Asians, and 40%–53% of Hispanics. Due to the important role of GSTM1 in the metabolism of carcinogens and drugs, this deletion variant in GSTM1 has been demonstrated to contribute to cancer susceptibility as well as to the prognosis of certain cancers. Previous studies have assessed risk based on the presence or absence of the GSTM1 gene, without determining the exact number of gene copies in individuals who possess GSTM1. Furthermore, potential functional promoter SNPs in the GSTM1 gene could influence the protein expression level and thus change the response of cells to exogenous carcinogens and drugs. However, investigation into these promoter genetic variants is still limited.

In this study, we report the results of the association between exact CNVs and promoter SNPs in GSTM1 with HNC REC/SPT. Our results indicate that this approach to assessing genetic variation within GSTM1 is a more superior prognostic biomarker than just determining the presence/absence of the gene and will aid in the identification of high-risk/poor-outcome individuals. Future replication studies in other independent cohorts are warranted to confirm these findings.

Methods

Human liver tissues and GST expression detection

A total of 111 human liver tissue samples were obtained from the US Cooperative Tissue Network (Birmingham, AL, USA) and liver cytosols were prepared by 100,000 g centrifugation of tissue homogenized in 20 mM Tris-hydrochloric acid (pH 7.8)-buffered 0.25 M sucrose containing 0.5 mM ethylenediaminetetraacetic acid, 20 µM butylated hydroxytoluene, and 0.1 mM dithiothreitol, as previously described. GSTs were determined after glutathione-agarose affinity chromatography and wide-pore high-performance liquid chromatography, as described by Coles and Kadlubar.

Study subjects

The study population has been described previously. Briefly, patients included in this study were stage I and II head and neck cases enrolled in the randomized Retinoid Head and Neck Second Primary Trial from 1991 to 1999 at MD Anderson Cancer Center, in which patients either received daily low dose (30 mg/day) of 13-cis-retinoic acid or placebo for 3 years. Patients must have remained cancer free for at least 16 weeks following surgery and/or radiation treatment to be enrolled in the trial. There were no recruitment restrictions on age, gender, or ethnicity. Before randomization, patients were given a structured questionnaire that elicited information on sociodemographic factors, tobacco-use history, alcohol consumption, and other exposures. Clinical data were obtained by medical chart review. The definitions of “second primary tumor” and “recurrence” following the Warren and Gates criteria were provided previously.

This study included 441 head and neck patients. Among these cases, 133 experienced recurrence of disease and/or an SPT, while 308 had single primary disease. All patients signed written informed consent before participation in the study and the study was approved by the Institution Review Board of the University of Texas MD Anderson Cancer Center. “Never smokers” were individuals who had smoked less than 100 cigarettes during their lifetime. “Former smokers” were individuals who had stopped smoking for at least a year at the time of enrollment.

Determination of GSTM1 copy number

The gene copy number of GSTM1 within the genome was determined with a TaqMan copy number assay (Applied Biosystems, Foster City, CA, USA), which used genomic DNA as a template and ran as a triplex TaqMan real-time polymerase chain reaction (PCR) with Rnase P as the reference.
reference gene, along with a well-characterized reference sample (Coriell Institute for Medical Research, Camden, NJ, USA) with two copies of GSTM1 as a calibrator. Real-time PCR data were analyzed by the comparative Ct method to calculate relative changes in the copy number of GSTM1. All CNV genotypes were determined without knowledge of the REC/SPT status of the subjects.

**Determination of genotypes in the promoter of GSTM1**

GSTM1 promoter fragments from each individual were amplified using the forward primer (GSTM1-PF: 5’-CAG GTT GGA CAT TGT TCT CGT G-3’) and reverse primer (GSTM1-PR: 5’-CAG CTG CTT CGC ACT TCC CT-3’) to produce a 1924 bp fragment. Genetic variants were identified by direct sequencing of the PCR products with a GenomeLab DTCS Quick Start Kit in a Beckman GeXP Genome Analyzer (both from Beckman Coulter, Fullerton, CA, USA). The sequencing primers were GSTM1-seq1 (5’-GGA GTT TCT TCA GAC TCA CAA T-3’), GSTM1-seq2 (5’-CCT GGG CCT TAA AGC ATG AC-3’), and GSTM1-seq3 (5’-CAC AGA CCA CAT TTC CTT TAC-3’). Genetic variants were identified using Codon-Code Aligner software (CodonCode, Dedham, MA, USA), a program for sequence assembly and mutation detection.

**Construction of reporter plasmids**

To verify whether identified SNPs influenced the transcriptional activity of GSTM1, we constructed eight reporter plasmids encompassing base pairs from –1687 to +87 of the human GSTM1 gene promoter. The primers used to amplify the fragment were 5’-GAC TAC GCG TTA CTG AAG AAC ACA CAT GG-3’ and 5’-GAA TAG ATC TGC GGA TGT CCC AGT ACC-3’, which contain Mlu I and Bgl II restriction sites (the underlined sequences), respectively. After amplification using LA Tq polymerase (NEB, Ipswich, MA, USA), the PCR product with the GSTM1 promoter was digested with Mlu I and Bgl II then inserted into a pGL3-basic vector (Promega, Madison, WI, USA) with the firefly luciferase gene as a reporter. The resulting construct was designated “pC-A-C” (with 498C→426A→339C) after sequence verification. This P-498C→426A→339C construct was subsequently used as a template to generate seven other constructs containing all possible haplotypes (P-498C→426A→339T, P-498C→426G→339C, P-498C→426G→339T, P-498G→426A→339C, P-498G→426A→339T, P-498G→426G→339C, P-498G→426G→339T) using the Quickchange Site-Directed Mutagenesis Kit (Agilent Technologies, Santa Clara, CA, USA). All constructs were confirmed by direct sequencing using the Beckman GeXP Genome Analyzer.

**Cell culture and luciferase assay**

The squamous cell carcinoma oral cavity cell line T409 was cultured in Dulbecco’s modified Eagle’s medium/F12 medium supplemented with 10% fetal bovine serum, 2 mmol/L L-glutamine, 100 U/mL penicillin, and 100 U/mL streptomycin (Invitrogen, Carlsbad, CA, USA) in a humidified, 5% CO2 incubator at 37°C. For transient transfection, 10 × 104 cells were plated in a 96-well plate and grown to 70%–80% confluence. Lipofectamine™ 2000 (Invitrogen) was used to transfect GSTM1-pGL3 basic constructs and control plasmid pRL-SV40 into cells according to the manufacturer’s protocol. Luciferase activity was measured using a TD20/20 luminometer (Turner Designs, Sunnyvale, CA, USA). The empty pGL3-basic vector was also transfected into cells as a control. Fold increase was calculated by defining the activity of empty pGL3 basic vector as 1. Differences were determined by t-test and a P value of <0.05 was considered significant.

**Statistical methods**

Multivariate Cox regression was used to estimate hazard ratios (HRs) associated with CNV, genotypes, or haplotypes along with 95% confidence intervals (CIs), while adjusting for confounding variables such as age, gender, smoking status, alcohol consumption, tumor site (larynx, oral cavity, and pharynx), stage (I or II), and randomization (13-cis-retinoic acid chemoprevention or placebo). Kaplan–Meier curves were constructed to compare event-free survival by CNV, genotypes, or haplotypes and logrank tests were performed to compare differences between survival curves. All analyses were performed using the Intercooled Stata statistical software package (v 10.0; Stata, College Station, TX, USA). All statistical tests were two sided and a P value of 0.05 was used as the criterion of statistical significance.

**Results**

GSTM1 protein expression

**GSTM1 genetic variants**

To verify the relationship between CNV and GSTM1 expression, we tested 111 human liver samples. Of these, 50 (45.0%) subjects were without GSTM1 expression and 61 subjects (55.0%) expressed GSTM1 with about 145-fold GSTM1 expression variation.

The results also showed that subjects with at least two copies have much higher GSTM1 expression than those with only one copy number (0.339 vs 0.118, P < 0.001) (Figure 1). This could partially account for GSTM1 expression variants among individuals, but among carriers with
one copy of $GSTM1$, there was still an 80-fold variation in protein expression. Thus, we conducted resequencing of a 1924-bp section of the 5' promoter region and found multiple genetic variants (Table 1). However, due to the very small sample size, there was no significant correlation between the genetic variants and $GSTM1$ protein levels.

**Genetic variation screening**

Through TaqMan $GSTM1$ CNV assay, we determined that 57% of the study subjects had the $GSTM1$ homozygous deletion, 24% had one copy of $GSTM1$, and 19% had at least two copies of $GSTM1$. Of those with the $GSTM1$ gene, we sequenced the full-length promoter region of $GSTM1$ and identified one insert variant and 14 SNPs (Table 1). Among these, the $-888A>T$ and $-341C>T$ were novel whereas the other SNPs had been previously deposited in the National Center for Biotechnology Information database (−1543 TTCT insertion [rs55791819], −1529C>G [rs36210087], −1490A>G [rs36209763], −1143A>G [rs36209754], −498C>G [rs412543], −486C>G [rs3815029], −471C>T [rs55791819], −426G>A [rs412302], −344C>T [rs4147561], −339C>T [rs4147562], −304G>A [rs28549287], −164 C>T [rs36208869]). The allele frequencies for these variants range from 0.07 to 0.47 (Table 1).

**CNV of $GSTM1$ and REC/SPT-free survival of HNC**

Regression analysis revealed significantly decreased REC/SPT in subjects with at least two copies of $GSTM1$ (odds ratio OR = 0.57; 95% CI = 0.35–0.95) but not in those with one copy of the gene (Table 2). Kaplan–Meier survival analysis revealed longer REC/SPT-free survival in patients with at least two copies of $GSTM1$. However, there was no significant difference between $GSTM1$ null subjects and one-copy carriers (Figure 2). SNP analysis showed the G allele of $-498C>G$, the G allele of $-426A>G$, and the T allele of $-339C>T$ SNPs were

---

**Table 1** List of genetic variants in the promoter of $GSTM1$ identified in Caucasians

<table>
<thead>
<tr>
<th>Genetic variants</th>
<th>Position*</th>
<th>rs number</th>
<th>MAF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. INS</td>
<td>−1542</td>
<td>NA</td>
<td>40.0</td>
</tr>
<tr>
<td>2. C/G</td>
<td>−1529</td>
<td>rs36210087</td>
<td>7.3</td>
</tr>
<tr>
<td>3. A/G</td>
<td>−1490</td>
<td>rs36209763</td>
<td>19.3</td>
</tr>
<tr>
<td>4. A/G</td>
<td>−1143</td>
<td>rs36209754</td>
<td>24.5</td>
</tr>
<tr>
<td>5. A/T</td>
<td>−888</td>
<td>NA</td>
<td>11.0</td>
</tr>
<tr>
<td>6. C/G</td>
<td>−498</td>
<td>rs412543</td>
<td>14.6</td>
</tr>
<tr>
<td>7. C/G</td>
<td>−486</td>
<td>rs3815029</td>
<td>9.9</td>
</tr>
<tr>
<td>8. C/T</td>
<td>−471</td>
<td>NA</td>
<td>12.8</td>
</tr>
<tr>
<td>9. A/G</td>
<td>−426</td>
<td>rs412302</td>
<td>26.7</td>
</tr>
<tr>
<td>10. C/T</td>
<td>−344</td>
<td>rs4147561</td>
<td>28.5</td>
</tr>
<tr>
<td>11. A/T</td>
<td>−343</td>
<td>rs4147562</td>
<td>25.8</td>
</tr>
<tr>
<td>12. C/T</td>
<td>−341</td>
<td>NA</td>
<td>15.9</td>
</tr>
<tr>
<td>13. C/T</td>
<td>−339</td>
<td>rs4147563</td>
<td>47.3</td>
</tr>
<tr>
<td>14. G/A</td>
<td>−304</td>
<td>rs28529287</td>
<td>18.5</td>
</tr>
<tr>
<td>15. T/C</td>
<td>−164</td>
<td>rs36208869</td>
<td>22.5</td>
</tr>
</tbody>
</table>

*Upstream of the ATG start site of the $GSTM1$ gene.

**Table 2** Copy number frequencies of $GSTM1$ and risk of SPT/REC in head and neck cancer

<table>
<thead>
<tr>
<th>CNV</th>
<th>No SPT/REC, N (%)</th>
<th>SPT/REC, N (%)</th>
<th>HR (95% CI)*</th>
<th>P*</th>
<th>Logrank test</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>167 (54.2)</td>
<td>85 (63.9)</td>
<td>0.73 (0.48–1.12)</td>
<td>0.15</td>
<td>0.07</td>
</tr>
<tr>
<td>1</td>
<td>76 (24.7)</td>
<td>29 (21.8)</td>
<td>0.57 (0.35–0.95)</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>2+</td>
<td>65 (21.1)</td>
<td>19 (14.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Notes: Adjusted for age, sex, race, smoking status, alcohol consumption, tumor site, tumor stage, and randomization arm; P values of multivariate Cox regression; P values of logrank test.

**Abbreviations:** CI, confidence interval; GSTM1, glutathione S-transferase Mu class 1; HR, hazard ratio; REC, recurrence; SPT, second primary tumor; CNV, copy number variant.
significantly associated with REC/SPT in HNC patients with HR 0.11 (0.02–0.85), 0.28 (0.11–0.74), and 2.02 (1.07–3.82), respectively (Table 3). Kaplan–Meier survival analysis also demonstrated that the −498G, −426G, and −339C alleles were significantly associated with longer REC/SPT-free survival times in HNC subjects (Figure 3A–C).

Haplotypes of GSTM1 promoter variants and REC/SPT-free survival of HNC

Three SNPs (−498C>G, −426A>G, and −339C>T) in the promoter of GSTM1 were independently associated with REC/SPT in HNC patients and were used for the construction of haplotypes. We evaluated the influence of different haplotypes on REC/SPT in our study population. The data showed that the patients with haplotype P−498G−426G−339C had decreased REC/SPT (HR = 0.09, 95% CI 0.01–0.71) compared with those with haplotype P−498C−426A−339T (Table 4).

There was no significant difference between other haplotypes and haplotype P−498C−426A−339T (Table 4). Kaplan–Meier survival estimates showed increased REC/SPT in haplotype P−498G−426G−339C group when compared with haplotype P−498C−426A−339T group (Figure 4).

**Effect of GSTM1 haplotypes on transcriptional activity**

To evaluate the influence of GSTM1 promoter SNPs on transcriptional activity, eight luciferase reporter constructs were generated. The sequences of these constructs encompassed the three SNPs (ie, −498G>C, −426G>A, and −339C>T) and were transiently transfected into T409 cells. As shown in Figure 5, reporter gene expression driven by haplotype P−498G−426G−339C of GSTM1 was 4.8-fold greater than with haplotype P−498C−426A−339T (P < 0.01). There were also statistically significant differences between other haplotypes and haplotype P−498C−426A−339T (Figure 5).

**Discussion**

This study explored the association between exact copy number of the GSTM1 gene and REC/SPT of HNC in a cohort study and provided compelling evidence of an association between HNC outcome and GSTM1 promoter SNPs/haplotype. One of the major findings is that the exact CNV can predict REC/SPT in a gene-dosage-dependent manner, with subjects having more than two copies of GSTM1 exhibiting the lowest risk of REC/SPT and longest REC/SPT-free survival. Another key finding is that three promoter SNPs (−498C>G, −426A>G, and −339C>T) were significantly associated with REC/SPT in HNC and the haplotype P−498G−426G−339C is associated with decreased risk of REC/SPT and improved REC/SPT-free survival. Moreover, it should be noted that all the significant associations remained significant after adjustment for cigarette

**Table 3** Allele frequencies of GSTM1 promoter SNPs and the risk of SPT/REC in head and neck cancer

<table>
<thead>
<tr>
<th>SNPs</th>
<th>No SPT/REC, N (%)</th>
<th>SPT/REC, N (%)</th>
<th>HR (95% CI)*</th>
<th>P*</th>
<th>Logrank test†</th>
</tr>
</thead>
<tbody>
<tr>
<td>−498C&gt;G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele C</td>
<td>99 (69.72)</td>
<td>43 (30.28)</td>
<td>1.00 (Ref)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele G</td>
<td>26 (96.30)</td>
<td>1 (3.70)</td>
<td>0.11 (0.02–0.85)</td>
<td>0.04</td>
<td>0.004</td>
</tr>
<tr>
<td>−426A&gt;G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele A</td>
<td>79 (67.52)</td>
<td>38 (32.48)</td>
<td>1.00 (Ref)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele G</td>
<td>46 (88.46)</td>
<td>6 (11.54)</td>
<td>0.28 (0.11–0.74)</td>
<td>0.01</td>
<td>0.002</td>
</tr>
<tr>
<td>−339C&gt;T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele C</td>
<td>82 (80.39)</td>
<td>20 (19.61)</td>
<td>1.00 (Ref)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele T</td>
<td>43 (64.18)</td>
<td>24 (35.82)</td>
<td>2.02 (1.07–3.82)</td>
<td>0.03</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Notes: *Adjusted for age, sex, race, smoking status, alcohol consumption, tumor site, tumor stage, and randomization arm; †P values of multivariate Cox regression; ‡P values of logrank test.

Abbreviations: CI, confidence interval; GSTM1, glutathione S-transferase Mu class 1; HR, hazard ratio; REC, recurrence; SNP, single nucleotide polymorphism; SPT, second primary tumor.
smoking, alcohol consumption, tumor site, tumor stage, and demographic factors, supporting exact copy number of the GSTM and GSTM1 promoter SNPs/haplotype as independent predictors of REC/SPT in HNC patients.

The incidence of HNC is increasing worldwide, and is associated with high mortality, especially in patients experiencing REC or a SPT.24 Previous studies have suggested that continued smoking, alcohol drinking, and tumor prognostic factors appear to be associated with the likelihood of SPT development.4,25,26 Recently, genetic variants in multiple cellular pathways have been shown to be independent predictors of REC/SPT in HNC patients.20–23 Long-term exposure to certain carcinogens contributes to REC/SPT of HNC, which is an unquestionably smoking-related cancer27–30 and is tightly related to the metabolism of carcinogens. As one of the major hepatic drug-metabolizing enzymes, GSTM1 plays an important role in the detoxification of carcinogens. At time of writing, at least 77 publications have reported on the association between the GSTM1 deletion variant and head and neck squamous cell carcinomas (HNSCC)11,31 and seven publications have examined the association between the GSTM1 deletion variant and outcomes of HNSCC.7,14,32,33 However, these studies have produced conflicting results concerning GSTM1 absence and risk and outcomes of HNSCC.14,30,32,34,35 In our study, we used exact copy number

---

**Table 4 GSTM1 haplotype frequencies and risk of SPT/REC in head and neck cancer**

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>No SPT/REC, N (%)</th>
<th>SPT/REC, N (%)</th>
<th>HR* (95% CI)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-G98C−426A−339T</td>
<td>39 (62.9)</td>
<td>23 (37.1)</td>
<td>1.00 (ref)</td>
<td></td>
</tr>
<tr>
<td>p-G98C−426A−339C</td>
<td>40 (72.7)</td>
<td>15 (27.3)</td>
<td>0.72 (0.37–1.42)</td>
<td>0.35</td>
</tr>
<tr>
<td>p-G98C−426G−339T</td>
<td>4 (80.0)</td>
<td>1 (20.0)</td>
<td>0.50 (0.02–10.54)</td>
<td>0.66</td>
</tr>
<tr>
<td>p-G98C−426G−339C</td>
<td>16 (80.0)</td>
<td>4 (20.0)</td>
<td>0.41 (0.14–1.47)</td>
<td>0.13</td>
</tr>
<tr>
<td>p-G98G−426G−339C</td>
<td>26 (96.3)</td>
<td>1 (3.7)</td>
<td>0.09 (0.01–0.69)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Notes: *Adjusted for age, sex, race, smoking status, alcohol consumption, tumor site, tumor stage, and randomization arm; **P** values of multivariate Cox regression.

Abbreviations: CI, confidence interval; GSTM1, glutathione S-transferase Mu class 1; HR, hazard ratio; REC, recurrence; SPT, second primary tumor.
of *GSTM1* instead of *GSTM1* presence/absence. We found that two or more copies of *GSTM1* conferred decreased risk of REC/SPT but one copy did not. In a prostate and bladder cancer study, researchers showed the exact copy number of *GSTM1* could predict the risk of bladder cancer, rather than just gene deletion. Our finding is also consistent with another study that showed a gene-dosage effect between copy number of *GSTM1* and enzyme activity.

Besides copy number, functional promoter genetic variants could influence gene expression levels by altering the binding ability of transcriptional factors to the gene promoter. In silico prediction modeling indicated multiple putative binding sites of transcriptional factors, such as AP2, GATA1, PEA3, and RXR/RAR, in the promoter region of *GSTM1*. Moreover, increased active transcriptional activity in the region from −600 to +34 of *GSTM1* after response to the *Myb* gene has been reported. In this study, we sequenced the promoter of *GSTM1* and found 14 genetic variants, including one insertion variant and 13 SNPs. Furthermore, we demonstrated that three SNPs (−498C>G, −426A>G, and −339C>T) are related to outcomes of HNC. To date, there is one other report related to SNPs in the promoter region of the *GSTM1* gene: in that study, Singh and colleagues selected three potentially functional SNPs and found −498C>G was associated with decreased risk of breast cancer. This is consistent with the present study in that −498G was related to reduced risk of REC and SPT and was also associated with increased REC/SPT-free survival in HNC. However, Yu et al reported that the −498G allele decreased gene transcription by 30%–40% by reducing the DNA-binding affinity of AP2 for the promoter region. This finding is inconsistent with the results of their reported case-control study. Since there are 14 common genetic variants in the promoter region of *GSTM1*, it is possible that *GSTM1* transcription is regulated by multiple transcription factors and could be influenced by several genetic variants in the promoter of *GSTM1*. In our HNC cohort study, in addition to −498C>G SNP, we also demonstrated that −426A>G and −339T>C were significantly associated with decreased REC/SPT and increased REC/SPT-free survival. This implies the complexity of *GSTM1* regulation by different transcriptional factors.

Since three promoter SNPs were independently associated with HNC outcomes, we then constructed haplotypes to examine the concerted effects of these variants. We found significant decreased REC/SPT and increased REC/SPT-free survival in patients with haplotype P-498G−426G−339C. Furthermore, in vitro assays demonstrated that haplotype P-498G−426G−339C has more transcriptional activity than other haplotypes. These data also support haplotype analysis as a better way of predicting the outcomes of HNC than examining single genetic variants.

Our study has limitations due to its relatively small sample size, so we were not able to assess the combined effects of CNV and promoter haplotypes on HNC outcomes; as such, examination of a larger study population is needed to clarify this relationship. However, this study is ongoing and can therefore be re-examined as more patients are accrued. The findings presented here should also be validated in other populations.

Despite these limitations, our findings support the effect of promoter SNPs and CNV in *GSTM1* on the outcomes of HNC. Further study is needed on the regulation of *GSTM1* expression and the influence of genetic variants—including CNV, SNP, or insertion—on the expression of *GSTM1* in human populations.

**Acknowledgments**

This work was supported by the Arkansas Biosciences Institute, the University of Arkansas for Medical Sciences’ Clinical Research Center, grant M01 RR14288. Fred Kadlubar passed away on December 4, 2010.

**Disclosure**

The authors declare no conflicts of interest in this work.

**References**


Pharmacogenomics and Personalized Medicine

Publish your work in this journal
Pharmacogenomics and Personalized Medicine is an international, peer-reviewed, open access journal characterizing the influence of genotype on pharmacology leading to the development of personalized treatment programs and individualized drug selection for improved safety, efficacy and sustainability. This journal is indexed on the American Chemical Society’s Chemical Abstracts Service (CAS). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: http://www.dovepress.com/pharmacogenomics-and-personalized-medicine-journal