Published in final edited form as:
*Ann Epidemiol*. 2009 August ; 19(8): 546–552. doi:10.1016/j.annepidem.2009.03.003.

# Copy number variants of *GSTM1* and *GSTT1* in relation to lung cancer risk in a prospective cohort study

**Tram Kim Lam**[1,2], **Ingo Ruczinski**[1], **Kathy Helzlsouer**[1,3,5], **Yin Yao Shugart**[1], **Kelly E. Li**[4], **Sandra Clipp**[1,5], **Paul T. Strickland**[6], and **Anthony J. Alberg**[1,5,7]

[1]Department of Epidemiology, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

[2]Cancer Prevention Fellowship Program, Office of Preventive Oncology, Division of Cancer Prevention, National Cancer Institute, National Institutes of Health, Bethesda, MD; Department of Biostatistics, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

[3]Mercy Medical Center, Baltimore, MD

[4]Applied Biosystems, Foster City, CA

[5]George W. Comstock Center for Public Health Research and Prevention, Hagerstown, MD

[6]Department of Environmental Health Sciences, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD

[7]Hollings Cancer Center and Department of Biostatistics, Bioinformatics, and Epidemiology, Medical University of South Carolina, Charleston SC

## Abstract

**PURPOSE—**Previous studies did not discriminate wild-type from hemizygous genotypes of *GSTM1* and *GSTT1*. In this study, we investigated wild-type, hemizygous deletion, and homozygous deletion genotypes of *GSTM1* and *GSTT1* and lung cancer risk.

**METHODS—**We conducted a nested case-control study of 143 primary incident lung cancer cases and 447 matched to cancer-free controls Genotyped data were obtained using a real-time PCR-based assay. Conditional logistic regression models were used to estimate odds ratios (ORs) and 95% Confidence Intervals (CIs).

**RESULTS—**Compared to *GSTM1* wild-type carriers, the relative odds of lung cancer increased from 1.49 (95% CI=0.66–3.40) to 1.80 (95% CI=0.81–4.02) for the hemizygous and homozygous deletion genotypes, respectively (p-trend=0.13). The strongest associations were seen among those who smoked <1 pack/day and had ≥deletion variant of *GSTM1* (OR=3.25; 95% CI=0.93–11.34; p-trend=0.07) whereas the reverse was observed for smokers who smoked ≥1 pack per day (OR=0.80; 95% CI=0.24–2.67; p-interaction=0.08). No clear associations were observed for *GSTT1* genotypes.

**CONCLUSIONS—**Risk of lung cancer increased as the number of deletion variants increased for *GSTM1*, though the associations were non-significant. Discriminating between the wild-type, hemizygous, and homozygous deletion *GSTM1* genotypes permitted a more precise characterization of the associations between *GSTM1* deletion variants and lung cancer.

**Correspondence and reprint requests to:** Anthony J. Alberg, Hollings Cancer Center, Medical University of South Carolina, PO Box 250955, Charleston, SC, 29425, 843-792-3426, 843-7924233 (fax), e-mail: E-mail: alberg@musc.edu.

### Keywords

Glutathione S-transferase; *GSTM1*; *GSTT1*; lung carcinoma; hemizygous deletion; homozygous deletion; copy number variants; metabolic genes

## INTRODUCTION

Lung cancer is the leading cause of cancer death worldwide (1). Cigarette smoking is responsible for approximately 85% of lung cancer deaths (2) and constituents of cigarette smoke such as polycyclic aromatic hydrocarbons (PAHs) are known carcinogens (3). Hydrophobic PAH-derived electrophiles bind to DNA to form PAH-DNA adducts, which may initiate carcinogenesis (4–6). Genes belonging to the *glutathione S-transferase (GST)* family detoxify tobacco-related carcinogens such as PAHs before the carcinogens damage DNA (7). *GSTM1* and *GSTT1* have been identified to have a deletion genotype (homozygous deletion) that results in a total absence of enzymatic activity. Because of their functional role in detoxifying tobacco-related carcinogens, variants of GSTs, such as *GSTM1* and *GSTT1*, may modulate lung cancer risk and contribute to individual susceptibility (7).

An extensive body of evidence has accumulated on the association between *GSTM1* and *GSTT1* and lung cancer risk comparing null-versus-present genotypes (8). Greater concentrations of PAH-DNA adducts have been found in lung tissues of *GSTM1* homozygous deletion (“null”, or 0/0) smokers compared to smokers with *GSTM1* present genotypes (wild-type or “+/+” plus hemizygous deletion or “ +/0” genotypes) (9). There is evidence of increased lung cancer risk among East Asians with *GSTM1* homozygous deletion, but not for Caucasians (meta-analysis of 98 studies) (10). Until recently genotyping assays were unable to discriminate between wild-type and hemizygous deletion genotypes (11). Hence, most of the evidence on this topic to date has not distinctly accounted for hemizygous *GSTM1* and *GSTT1* genotypes. Consequently, previous investigations often compared the risks associated with the homozygous deletion genotype to a referent “present” genotype that was comprised of both wild-type and hemizygous deletion genotypes. Data suggest that distinguishing the wild-type from the hemizygous genotype is important, as studies using new real-time PCR-based assays that discriminate between the wild-type, hemizygous deletion, and homozygous deletion *GSTM1* and *GSTT1* genotypes show these genotypes to exhibit a high, intermediate, and absence of enzymatic activity, respectively (12–15). This raises the possibility that previously published results may tend to underestimate the true associations of *GSTM1* and *GSTT1* in relation to lung cancer risk because this misclassification would bias results toward the null (16). Support for this hypothesis has been observed for other malignancies (17–19). The present study was carried out to test whether the lung cancer risk varies according the number of functional alleles of *GSTM1* and *GSTT1* using a case-control study nested within a community-based cohort in Washington County, Maryland.

## METHODS

### Study population

Established in 1989, the CLUE II cohort was named for its campaign slogan, “*Give Us a Clue to Cancer and Heart Disease*.” The details of the establishment of the cohort have been reported elsewhere (20,21). Briefly, from May through October 1989, 32,897 residents of Washington County, Maryland agreed to participate in CLUE II. Of these individuals, 25,081 adults (>18 year old) provided a Washington County address and are covered by the county cancer registry. Starting in 1996, the CLUE II cohort members received periodic follow-up questionnaires.

At baseline, basic demographic characteristics, smoking status, and number of cigarette smoked per day were obtained using a brief questionnaire. At that time, participants also provided blood samples (20 ml) drawn into a 20 mL Vacutainer tube containing heparin and immediately refrigerated until centrifugation. Centrifugation usually took place within 6 hours and never exceeded twenty-four hours. Once centrifuged, aliquots of plasma, red blood cells, and buffy coats were separated and stored at-70° C in a specimen bank.

Ascertainment of lung cancer cases was achieved through linkage with the Washington County Cancer Registry, the Maryland State Cancer Registry, and death certificates. The Washington County Cancer Registry receives its data primarily from the county's only general hospital, Washington County Hospital. CLUE II cohort members were also linked to the Maryland State Cancer Registry, which was established in 1993 and has a mandatory cancer reporting policy.

## Case and control selection

The present study is comprised of 143 incident lung cancer cases that occurred from 1990 to October, 2005 and 447 matched cancer-free controls with genotype data for at least one *GSTM1* or *GSTT1*. With the exception of non-melanoma skin cancer or cervical carcinoma *in situ*, cases were first-time cancer diagnosed with primary lung cancer (*International Classification of Diseases, Eighth Revision [ICD-8]* for cases diagnosed before 1992 code 162; *ICD-9* code 162 for cases diagnosed from 1992–2000, and *ICD-10* codes C33-C34 for cases diagnosed from 2000-present).

For each case, we selected up to four controls. At least one control was matched to each case (n=17 case-control sets with one control) and the average number of controls per case was three. Eligibility criteria for control selection were: (1) completion of the baseline questionnaire; (2) no prior history of cancer except for non-melanoma skin cancer or cervical cancer *in situ*; (3) cancer-free and known to be alive at the time of case diagnosis.

Controls were individually matched to cases on the following variables: gender, age (±5 years), and smoking status (never, former, or current smokers). For former and current smokers, cases and controls were further matched on the number of cigarettes smoked per day as follows. Ever smokers were categorized into three smoking groups: ≤ 19, 20 to 29, and ≥ 30 cigarettes per day. For individuals who smoked <30 cigarettes per day, cases and controls were matched within ±5 cigarettes smoked per day. Those who smoked between 30–45 cigarettes per day were matched within ±10 and the heaviest smokers (≥45) were matched within ±20 cigarettes. The 1996 follow-up questionnaire was relevant to this study as it provided additional detail on pack-years of smoking on cases and controls. Among cases who provided information on the 1996 follow-up questionnaire about pack-years of smoking, we matched controls to within ±5 pack-years of cigarette smoking.

## *GSTM1* and *GSTT1* genotyping

DNA was extracted from buffy coat, frozen at −70°C, using an alkaline lysis method (22). *GSTM1* and *GSTT1* genotyping was performed by Applied Biosystems (Foster City, California; www.appliedbiosystems.com) using TaqMan® Gene Copy Number Assays (PN4331182). The assay consisted of two primers and a FAM™-MGB probe in 20X formulation and measurements were made in real time. Primers and probes were designed from genomic sequence (hg18/Build 36) using Applied Biosystems proprietary software. For quality control, each assay was run as a duplex TaqMan real-time PCR reaction, one containing a FAM dye-based assay for the targeted gene and a VIC dye-based assay for the reference gene. An additional, a known sample was inserted randomly in each batch which was blinded to the technician performing the assay.

All assays were conducted in a 96-well plate (MicroAmp™ Optical 384-Well Reaction Plate). Each well contained FAM-labeled TaqMan probe for either *GSTM1* or *GSTT1* and VIC-labeled TaqMan probe for the reference gene. PCR was performed in a reaction mixture containing the following: 2x TaqMan® Universal PCR Master Mix; both *GSTM1* primer and probe or *GSTT1* primer and probe; DNase-free water; and genomic DNA sample. Real-time data analysis was performed using Sequence Detection Software v2.1 (Absolute Quantification and Copy Number Macro, www.allgenes.com). Relative quantity is determined by the Ct $((\text{FAM Ct} - \text{VIC Ct})_{\text{sample}} - (\text{FAM Ct} - \text{VIC Ct})_{\text{calibrator}})$ method, where a reference sample or calibrator known to have two copies of the test sequence is used as the basis for comparative results. The gene copy number is two times the relative quantity (23).

### Statistical analysis

Hardy-Weinberg Equilibrium (HWE) was assessed based on the distribution of the *GSTM1* and *GSTT1* genotypes among the controls using SIBPAIR version 0.99.0 (http://www.qimr.edu.au/davidD/davidd.html). Unless otherwise indicated, conditional logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI) for lung cancer risk comparing the *GSTM1* and *GSTT1* wild-type to the hemizygous deletion and the homozygous deletion genotypes, respectively.

All estimates were adjusted for age (continuous) and number of cigarettes smoked per day (continuous) to account for possible residual confounding that could persist after matching. Additional analyses were stratified by gender, above-versus-below the mean age of the controls, smoking status (never, former, and current), and above-versus-below 20 cigarettes smoked per day (median among the controls). The likelihood ratio test was used to test for dose-response trends across the three genotypes *GST* genotypes by fitting a single genotype variable as a single categorical variable. To assess how the results of analyses using the refined genotyping compared to the traditional dichotomous null-versus-present genotyping, analyses were also performed comparing the homozygous deletion genotype to a referent group comprised of the combined hemizygous deletion (+/0) plus wild-type genotypes (+/+). A two-tailed *p-value* of $< 0.05$ was considered to be statistically significant. All analyses were performed using STATA version 9.1

## RESULTS

Cases and controls were matched on age, gender and smoking status and did not differ significantly with regard to other selected demographic characteristics (Table I). On average, the study population was 62 years of age at baseline. As expected, never smokers comprised only a small percentage (~8%) of lung cancer cases.

Among the controls, the frequencies observed for *GSTM1* wild-type, hemizygous deletion, and homozygous deletion genotypes were 9%, 40%, and 50% respectively (Table II). For *GSTT1*, the distribution was 24%, 51%, and 20% for wild-type, hemizygous deletion, and homozygous deletion genotypes, respectively. These frequencies of *GSTM1* and *GSTT1* homozygous deletion genotype seen in the controls were similar to those previously observed for Caucasians (8). The tests for HWE showed no deviation (for *GSTM1*: p-value = 0.80 and for *GSTT1*: p-value=0.11).

Compared to those with the *GSTM1* wild-type genotype, the risks of lung cancer were 1.49 (95% CI =0.66–3.40) and 1.80 (95% CI = 0.81–4.02) for those with the hemizygous deletion and homozygous deletion genotypes, respectively (p- trend = 0.13, Table III). For *GSTT1*, compared to the wild-type genotype, the risks of lung cancer were 1.17 (95% CI = 0.71–1.92) and 1.06 (95% CI= 0.58–1.95) for those with the hemizygous and homozygous deletion genotypes, respectively (p-= 0.83). None of these associations were statistically significant.

When stratified by smoking status, the odds ratios for lung cancer among those with the hemizygous and homozygous deletion *GSTM1* genotypes ranged between 1.40 and 1.99 in both former and current smokers (Table III). Among those who smoked <=20 cigarettes per day equals to 1 pack or less per day, compared to the wild-type genotype the odds ratios were 3.06 (95% CI= 0.83–11.28; p-value= 0.09) for the hemizygous deletion genotype and 3.35 (95% CI= 0.94–11.86; p-value= 0.06) for the homozygous deletion genotype (p-trend= 0.12) (Table III). In contrast, the odds ratios were slightly in the protective direction for smokers who smoked > 1pack per day. The p-value for the test for interaction by smoking intensity was 0.07.

The smoking-stratified results for *GSTT1* were opposite those seen for *GSTM1*. Specifically, no associations were seen in ever smokers who smoked ≤1 pack per day, whereas among heavier smokers the associations for the hemizygous deletion and homozygous deletion *GSTT1* genotypes were in the direction of increased risk.

To illustrate the potential value of the refined classification of *GST* genotypes that explicitly accounts for hemizygotes, we reanalyzed our data using the traditional null-versus-present comparison in which the referent group included the hemizygous deletion *GSTM1* genotype. When the data were re-classified to conform to the traditional null-versus-present genotype, the risks associated with the *GSTM1* genotype were always attenuated toward the null compared to the more refined classifications we presented in our primary analyses (Null vs present comparison, Table III). This was particularly true among lighter smokers among whom the odds ratios diminished from 3.35 to 1.31 for the *GSTM1* homozygous deletion variant.

## DISCUSSION

*GSTM1* and *GSTT1* have been thoroughly investigated in relation to lung cancer risk because of their critical role in inactivating tobacco-related carcinogens (24). The value of the present study was the use of a real-time PCR-based assay for genotyping that distinguishes between the three distinct genotypes of wild-type, hemizygous deletion, and homozygous deletion, whereas most previous studies could only distinguished the homozygous deletion, “null” genotype, versus the “present” genotype consisting of the wild-type and hemizygous deletion genotypes. Based on the *GSTM1* null-versus-present comparisons used predominantly in the past, overall weak or no associations have been observed (25–33). Notably, when we reanalyzed our *GSTM1* data using this traditional genotyping classification, the associations were attenuated. This is consistent with misclassification of those lacking a functional allele. The results were not statistically significant, but the overall pattern of associations suggested the risk of lung cancer increased as the copy number of *GSTM1* deletion variants increased from zero (wild-type) to one (hemizygous deletion) to two (homozygous deletion). These results are consistent with previous observation that hemizygous and wild-type genotypes are associated with function difference in enzymatic activity (15).

The only previous study we are aware of to report on the associations between *GSTM1* and *GSTT1* genotypes and lung cancer risk that classified hemizygous deletion separately from wild-type observed no difference between the hemizygous and homozygous null genotypes of *GSTM1* and *GSTT1* when compared to the referent wild-type genotype (34). Our results are in agreement with the previous lack of associations observed for *GSTT1*, and thus suggest caution in interpreting our results for *GSTM1*. Compared to the previous study, the careful matching of cases and controls on smoking exposures is a notable strength of our study design. This is primarily because cigarette smoking is the predominant cause of lung cancer such that even the residual effects of smoking could overwhelm smaller risks associated with genetic factors.

Although not statistically significant, our smoking-stratified results suggested that lung cancer risk associated with *GSTM1* deletions was more pronounced among smokers who smoked ≤ 1 pack of cigarettes per day, whereas the converse was observed for heavy smokers. The lack of association among heavier smokers could possibly be due to the carcinogen exposure levels being so high that they overwhelm the effects of the *GSTM1* deletion genotype. Our findings of increased risk for light smokers are consistent with other findings from case-control studies on *GSTM1*-smoking interaction in relation to lung cancer (35–37), suggesting that *GSTM1* may have a more discernable influence in the milieu of lower carcinogenic exposure. The findings of our study also may be due to chance.

The present study benefits from the prospective study design that avoids methodological limitations of retrospective studies such as selection and recall bias. In this setting, the relevance of recall bias primarily pertains to the measurement of cigarette smoking. Blood samples were collected at baseline and before onset of lung cancer diagnosis. This minimizes the possibility of survival bias that might be attributable to *GSTM1* variants (38,39).

An important limitation of this study is that it lacked adequate statistical precision to detect statistical significance that may be associated with *GST* genotypes. We focused on only a few genes, but a pathway-based approach, in which the concerted risk of multiple at-risk variants of metabolic genes are investigated, would provide a more comprehensive characterization of variant genes and lung carcinogenesis (40). For example, a deficiency in one or two genes may be compensated by other genes within the *GST* pathway (40,41).

In summary, the results of this nested case-control study provide limited support for the hypothesis that accounting for the hemizygous deletion *GSTM1* genotype strengthens the association between *GSTM1* genotype and lung cancer risk compared to previous evidence based on present-versus-null genotype comparisons. The results were not statistically significant, but these hypothesis-generating findings imply additional investigations with genotyping for hemizygote status will enhance the resolution of our understanding of this question.

## ABBREVIATIONS

*GSTs*, Glutathione *S*-transferases; *GSTM1*, *GST Mu 1*; *GSTT1*, *GST Theta 1*; PCR, Polymerase Chain Reaction; PAH, polycyclic aromatic hydrocarbons; OR, Odds ratio; 95%CI, 95% confidence interval; ICD, International Classification of Diseases; HWE, Hardy-Weinberg Equilibrium.

## ACKNOWLEDGEMENTS

This work was supported by the Institutional Research Epidemiology Fellowship (T32CA009314), National Institute of Aging (5U01AG018033) and National Cancer Institute (CA105069). The authors express their appreciation to the participants of the CLUE II cohort and thank the staff at the George W. Comstock Center for Public Health Research and Prevention for their dedication and contributions to the study.

## REFERENCES

1. Ferlay, J.; Bray, F.; Pisani, P.; Parkin, DM. IARC CancerBase No. 5, version 2.0. IARC Press; 2004. GLOBOCAN 2002: Cancer Incidence, Mortality and Prevalence Worldwide. Edition Lyon: IARC *Press*
2. The Health Consequences of Smoking. Washington DC: Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2004. A Report of the Surgeon General. Edition
3. Hecht SS. Tobacco Smoke Carcinogens and Lung Cancer. Journal of the National Cancer Institute 1999;91:1194–1210. [PubMed: 10413421]

4. Phillips DH. DNA adducts as markers of exposure and risk. Mutat Res 2005;577:284–292. [PubMed: 15922369]
5. Phillips DH. Smoking-related DNA and protein adducts in human tissues. Carcinogenesis 2002;23:1979–2004. [PubMed: 12507921]
6. Wiencke JK. DNA adduct burden and tobacco carcinogenesis. Oncogene 2002;21:7376–7391. [PubMed: 12379880]
7. Ketterer B. A bird's eye view of the glutathione transferase field. Chem Biol Interact 2001;138:27–42. [PubMed: 11640913]
8. Ye Z, Song H, Higgins JP, Pharoah P, Danesh J. Five glutathione s-transferase gene variants in 23,452 cases of lung cancer and 30,397 controls: meta-analysis of 130 studies. PLoS Med 2006;3:e91. [PubMed: 16509765]
9. Kato S, Bowman ED, Harrington AM, Blomeke B, Shields PG. Human lung carcinogen-DNA adduct levels mediated by genetic polymorphisms in vivo. J Natl Cancer Inst 1995;87:902–907. [PubMed: 7666479]
10. Carlsten C, Sagoo GS, Frodsham AJ, Burke W, Higgins JP. Glutathione S-transferase M1 (GSTM1) polymorphisms and lung cancer: a literature-based systematic HuGE review and meta-analysis. Am J Epidemiol 2008;167:759–774. [PubMed: 18270371]
11. Parl FF. Glutathione S-transferase genotypes and cancer risk. Cancer Lett 2005;221:123–129. [PubMed: 15808397]
12. Sprenger R, Schlagenhaufer R, Kerb R, Bruhn C, Brockmoller J, Roots I, et al. Characterization of the glutathione S-transferase GSTT1 deletion: discrimination of all genotypes by polymerase chain reaction indicates a trimodular genotype-phenotype correlation. Pharmacogenetics 2000;10:557–565. [PubMed: 10975610]
13. Girault I, Lidereau R, Bieche I. Trimodal GSTT1 and GSTM1 genotyping assay by real-time PCR. Int J Biol Markers 2005;20:81–86. [PubMed: 16011037]
14. Roodi N, Dupont WD, Moore JH, Parl FF. Association of homozygous wild-type glutathione S-transferase M1 genotype with increased breast cancer risk. Cancer Res 2004;64:1233–1236. [PubMed: 14973116]
15. Bediaga NG, Alfonso-Sanchez MA, de Renobales M, Rocandio AM, Arroyo M, de Pancorbo MM. GSTT1 and GSTM1 gene copy number analysis in paraffin-embedded tissue using quantitative real-time PCR. Anal Biochem 2008;378:221–223. [PubMed: 18440293]
16. Brasch-Andersen C, Christiansen L, Tan Q, Haagerup A, Vestbo J, Kruse TA. Possible gene dosage effect of glutathione-S-transferases on atopic asthma: using real-time PCR for quantification of GSTM1 and GSTT1 gene copy numbers. Hum Mutat 2004;24:208–214. [PubMed: 15300848]
17. Moore LE, Huang WY, Chatterjee N, Gunter M, Chanock S, Yeager M, et al. GSTM1, GSTT1, and GSTP1 polymorphisms and risk of advanced colorectal adenoma. Cancer Epidemiol Biomarkers Prev 2005;14:1823–1827. [PubMed: 16030123]
18. Garcia-Closas M, Malats N, Silverman D, Dosemeci M, Kogevinas M, Hein DW, et al. NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. Lancet 2005;366:649–659. [PubMed: 16112301]
19. Christiansen L, Brasch-Andersen C, Bathum L, Kruse TA, Christensen K. A longitudinal study of the effect of GSTT1 and GSTM1 gene copy number on survival. Mech Ageing Dev 2006;127:597–599. [PubMed: 16574194]
20. Helzlsouer KJ, Huang HY, Strickland PT, Hoffman S, Alberg AJ, Comstock GW, et al. Association between CYP17 polymorphisms and the development of breast cancer. Cancer Epidemiol Biomarkers Prev 1998;7:945–949. [PubMed: 9796641]
21. Hoffman SC, Burke AE, Helzlsouer KJ, Comstock GW. Controlled trial of the effect of length, incentives, and follow-up techniques on response to a mailed questionnaire. Am J Epidemiol 1998;148:1007–1011. [PubMed: 9829873]
22. Klintschar M, Neuhuber F. Evaluation of an alkaline lysis method for the extraction of DNA from whole blood and forensic stains for STR analysis. J Forensic Sci 2000;45:669–673. [PubMed: 10855975]
23. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402–408. [PubMed: 11846609]

24. Hayes JD, Strange RC. Glutathione S-transferase polymorphisms and their biological consequences. Pharmacology 2000;61:154–166. [PubMed: 10971201]
25. Dialyna IA, Miyakis S, Georgatou N, Spandidos DA. Genetic polymorphisms of CYP1A1, GSTM1 and GSTT1 genes and lung cancer risk. Oncol Rep 2003;10:1829–1835. [PubMed: 14534704]
26. Gao Y, Zhang Q. Polymorphisms of the GSTM1 and CYP2D6 genes associated with susceptibility to lung cancer in Chinese. Mutat Res 1999;444:441–449. [PubMed: 10521684]
27. Gsur A, Haidinger G, Hollaus P, Herbacek I, Madersbacher S, Trieb K, et al. Genetic polymorphisms of CYP1A1 and GSTM1 and lung cancer risk. Anticancer Res 2001;21:2237–2242. [PubMed: 11501853]
28. Kiyohara C, Yamamura KI, Nakanishi Y, Takayama K, Hara N. Polymorphism in GSTM1, GSTT1, and GSTP1 and Susceptibility to Lung Cancer in a Japanese Population. Asian Pac J Cancer Prev 2000;1:293–298. [PubMed: 12716303]
29. Lewis SJ, Cherry NM, Niven RM, Barber PV, Povey AC. GSTM1, GSTT1 and GSTP1 polymorphisms and lung cancer risk. Cancer Lett 2002;180:165–171. [PubMed: 12175548]
30. McWilliams JE, Sanderson BJ, Harris EL, Richert-Boe KE, Henner WD. Glutathione S-transferase M1 (GSTM1) deficiency and lung cancer risk. Cancer Epidemiol Biomarkers Prev 1995;4:589–594. [PubMed: 8547824]
31. Miller DP, Liu G, De Vivo I, Lynch TJ, Wain JC, Su L, et al. Combinations of the variant genotypes of GSTP1, GSTM1, and p53 are associated with an increased lung cancer risk. Cancer Res 2002;62:2819–2823. [PubMed: 12019159]
32. Nazar-Stewart V, Vaughan TL, Stapleton P, Van Loo J, Nicol-Blades B, Eaton DL. A population-based study of glutathione S-transferase M1, T1 and P1 genotypes and risk for lung cancer. Lung Cancer 2003;40:247–258. [PubMed: 12781423]
33. Schneider J, Bernges U, Philipp M, Woitowitz HJ. GSTM1, GSTT1, and GSTP1 polymorphism and lung cancer risk in relation to tobacco smoking. Cancer Lett 2004;208:65–74. [PubMed: 15105047]
34. Sorensen M, Raaschou-Nielsen O, Brasch-Andersen C, Tjonneland A, Overvad K, Autrup H. Interactions between GSTM1, GSTT1 and GSTP1 polymorphisms and smoking and intake of fruit and vegetables in relation to lung cancer. Lung Cancer 2007;55:137–144. [PubMed: 17123660]Epub 2006 Nov 2022
35. Cote ML, Kardia SL, Wenzlaff AS, Land SJ, Schwartz AG. Combinations of glutathione S-transferase genotypes and risk of early-onset lung cancer in Caucasians and African Americans: a population-based study. Carcinogenesis 2005;26:811–819. [PubMed: 15661806]
36. Alexandrie AK, Nyberg F, Warholm M, Rannug A. Influence of CYP1A1, GSTM1, GSTT1, and NQO1 genotypes and cumulative smoking dose on lung cancer risk in a Swedish population. Cancer Epidemiol Biomarkers Prev 2004;13:908–914. [PubMed: 15184245]
37. London SJ, Daly AK, Cooper J, Navidi WC, Carpenter CL, Idle JR. Polymorphism of glutathione S-transferase M1 and lung cancer risk among African-Americans and Caucasians in Los Angeles County, California. J Natl Cancer Inst 1995;87:1246–1253. [PubMed: 7563171]
38. London SJ, Yuan JM, Chung FL, Gao YT, Coetzee GA, Ross RK, et al. Isothiocyanates, glutathione S-transferase M1 and T1 polymorphisms, and lung-cancer risk: a prospective study of men in Shanghai, China. Lancet 2000;356:724–729. [PubMed: 11085692]
39. Sweeney C, Nazar-Stewart V, Stapleton PL, Eaton DL, Vaughan TL. Glutathione S-transferase M1, T1, and P1 polymorphisms and survival among lung cancer patients. Cancer Epidemiol Biomarkers Prev 2003;12:527–533. [PubMed: 12814998]
40. Saarikoski ST, Voho A, Reinikainen M, Anttila S, Karjalainen A, Malaveille C, et al. Combined effect of polymorphic GST genes on individual susceptibility to lung cancer. Int J Cancer 1998;77:516–521. [PubMed: 9679751]
41. Stranger B, Forrest M, Dunning. Relative Impact of Nucleotide and Copy Number Variation on Gene Expression Phenotypes. Science 2007;315:848–853. [PubMed: 17289997]

**Table I**

Baseline characteristics of lung cancer cases and controls, Washington County, MD (1989–2005)

| | Cases (n=143) | Controls (n=447) | p-value [a] |
|---|---|---|---|
| **Characteristic Age at base line, yrs (%)** | | | |
| Mean (±SD) | 61.67 (±8.49) | 61.46 (±8.36) | 0.80 |
| <45 (%) | 2.80 | 2.68 | |
| 45–64 (%) | 58.04 | 58.61 | |
| 65+ (%) | 39.16 | 38.70 | * |
| **Gender (%)** | | | |
| Female | 53.15 | 54.36 | * |
| **Marital Status (%)**[b] | | | |
| Single | 2.80 | 2.01 | |
| Married | 72.73 | 74.50 | |
| Other | 23.78 | 22.82 | 0.94 |
| **Education (%)**[c] | | | |
| < 12 years | 32.87 | 30.87 | |
| High school graduate | 44.96 | 46.31 | |
| Beyond high school | 22.38 | 22.82 | 0.34 |
| **Body mass index in kg/m$^2$** | | | |
| Mean (±SD) | 25.61 ± 4.01 | 25.81 ± 3.96 | 0.60 |
| <24.9 (%) | 47.55 | 44.30 | |
| 25.0–29.9 (%) | 40.56 | 41.39 | |
| 30+ (%) | 11.89 | 14.32 | 0.69 |
| **Family history of cancer** | | | |
| No | 58.74 | 50.78 | |
| Yes | 41.26 | 49.22 | 0.10 |
| **Cigarette smoking status (%)** | | | |
| Never (n=60) | 8.39 | 10.74 | |
| Former (n=270) | 44.06 | 46.31 | |
| Current (n=260) | 47.55 | 42.95 | * |
| **Cigarettes smoked per day (CPD)** | | | |
| All (Mean ±SD) † | 25.13 ±13.72 | 23.93 ±12.27 | 0.35 |

[a] p-value: Pearson $\chi^2$ or ttest for means

* Matching variables

* Matching variables

[b] Missing data for 1 case and 3 controls

[c] Missing data for 1 case

* Matching variables

† Ever smokers only

| | Cases (n=143) | Controls (n=447) | p-value [a] |
|---|---|---|---|
| *Former smokers* | | | |
| Mean (±SD) | 26.95 ±14.32 | 25.86 ±12.98 | 0.57 |
| ≤10 CPD (%) | 19.05 | 18.84 | |
| 11–39 CPD (%) | 39.68 | 42.51 | |
| 40+ CPD (%) | 41.27 | 38.65 | * |
| *Current smokers* | | | |
| Mean (±SD) | 23.44 ±13.02 | 21.87 ±11.11 | 0.34 |
| ≤10 | 27.94 | 27.60 | |
| 11–39 | 42.65 | 44.27 | |
| 40+ | 29.41 | 28.13 | * |

[*] Matching variables

[a] p-value: Pearson $\chi^2$ or ttest for means

[b] Missing data for 1 case and 3 controls

[c] Missing data for 1 case

[†] Ever smokers only

[*] Matching variables

[*] Matching variables

**Table II**

Genotype distributions of *GSTM1* and *GSTT1* for selected characteristics separated by cases and controls

| | Cases (n=143)* | | | Controls (n=447)* | | |
|---|---|---|---|---|---|---|
| | **All (%)** | **Female (%)** | **Male (%)** | **All (%)** | **Female (%)** | **Male (%)** |
| GSTM1‡ | | | | | | |
| +/+ | 9 (6.29) | 3 (3.95) | 6 (8.96) | 40 (8.95) | 14 (5.76) | 26 (12.75) |
| +/0 | 53 (37.06) | 31 (41.79) | 22 (32.84) | 179 (40.04) | 99 (40.74) | 80 (39.22) |
| 0/0 | 75 (52.45) | 40 (52.63) | 35 (52.24) | 222 (49.66) | 127 (52.26) | 95 (46.57) |
| Missing | 6 (4.20) | 2 (2.63) | 4 (5.97) | 6 (1.34) | 3 (1.23) | 3 (1.47) |
| GSTT1† | | | | | | |
| +/+ | 31 (21.68) | 15 (19.74) | 16 (23.88) | 108 (24,16) | 63 (25.93) | 45 (22.06) |
| +/0 | 75 (52.45) | 37 (48.68) | 38 (56.72) | 227 (50.78) | 122 (50.21) | 105 (51.47) |
| 0/0 | 29 (20.28) | 19 (25.00) | 10 (14.93) | 87 (19.46) | 44 (18.11) | 43 (21.08) |
| Missing | 8 (5.59) | 5 (6.58) | 3 (4.48) | 25 (5.59) | 14 (5.76) | 11 (5.39) |

*values in columns = number of cases or controls (%); % do not add up to 100% due to missing data

‡six cases and six controls had no *GSTM1* genotype data

†eight cases and 25 controls had no *GSTT1* genotype data

Genotypes: (+/+): wild-type genotype;(+/**0**): hemizygous deletion; (**0/0**): homozygous deletion

[*] values in columns = number of cases or controls (%); % do not add up to 100% due to missing data

[*] values in columns = number of cases or controls (%); % do not add up to 100% due to missing data

[‡] six cases and six controls had no *GSTM1* genotype data

[†] eight cases and 25 controls had no *GSTT1* genotype data

**Table III**

The relative odds of developing lung cancer according to *GSTM1* and *GSTT1* genotypes, total and stratified by cigarette smoking characteristics, Washington County, MD (1989–2005)

| | GSTM1 | | | | GSTT1 | | | |
|---|---|---|---|---|---|---|---|---|
| | **Genotype‡** | **Cases/Controls** | **OR (95% CI)[a]** | **p-value** | **Genotype‡** | **Cases/Controls** | **OR (95% CI)[a]** | **p-value** |
| **All subjects** | | **(137/441)** | | | | **(135/422)** | | |
| | +/+ | 9/40 | 1.00 | | +/+ | 31/108 | 1.00 | |
| | 0/0 | 75/222 | 1.80 (0.81–4.02) | 0.13[e] | 0/0 | 29/87 | 1.06 (0.58–1.95) | 0.83[e] |
| | +/0 & 0/0 | 128/401 | 1.67 (0.76–3.65) | 0.20 | +/0 & 0/0 | 104/314 | 1.14 (0.71–1.83) | 0.60 |
| | Null vs present[d] | | 1.29 (0.87–1.92) | | Null vs present[d] | | 0.95 (0.58–1.56) | |
| **Smoking status** | | | | | | | | |
| Never smokers[b] | | | | | | | | |
| | +/+ | 1/4 | 1.00 | | +/+ | 3/14 | 1.00 | |
| | +/0 | 5/21 | 1.00 (0.09–11.52) | | +/0 | 7/23 | 1.3 (0.3–5.8) | 0.63 |
| | 0/0 | 6/23 | 1.00 (0.09–10.99) | - | 0/0 | 0/8 | - | |
| | +/0 or 0/0 | 11/44 | 1.0 (0.10–10.07) | - | +/0 & 0/0 | 7/31 | 1.00 (0.2–4.5) | 0.96 |
| | Null vs present[δ] | | | 1.00 (0.26–3.89) | Null vs present[δ] | | - | |
| Former smokers | | | | | | | | |
| | +/+ | 4/22 | 1.0 | | +/+ | 13/53 | 1.00 | |
| | +/0 | 25/84 | 1.92 (0.57–6.49) | | +/0 | 40/106 | 1.65 (0.79–3.45) | |
| | 0/0 | 30/97 | 1.99 (0.63–6.24) | 0.34[e] | 0/0 | 8/36 | 0.95 (0.36–2.54) | 0.91[e] |
| | +/0 & 0/0 | 55/181 | 1.98 (0.64–6.10) | 0.24 | +/0 & 0/0 | 48/142 | 1.45 (0.71–2.96) | 0.30 |
| | Null vs present[δ] | | 1.20 (0.66–2.17) | | Null vs present[δ] | | 0.67 (0.30–1.51) | |
| Current smokers | | | | | | | | |
| | +/+ | 4/14 | 1.00 | | +/+ | 15/41 | 1.00 | |
| | +/0 | 23/74 | 1.40 (0.38–5.18) | | +/0 | 28/98 | 0.75 (0.34–1.64) | |

‡(+/+): wild-type genotype; (0/+): hemizygous deletion; (0/0): homozygous deletion

| | GSTM1 | | | | GSTT1 | | | |
|---|---|---|---|---|---|---|---|---|
| | **Genotype‡** | **Cases/Controls** | **OR (95% CI)[a]** | **p-value** | **Genotype‡** | **Cases/Controls** | **OR (95% CI)[a]** | **p-value** |
| | 0/0 | 39/102 | 1.89 (0.51–6.99) | 0.22[e] | 0/0 | 21/43 | 1.10 (0.47–2.59) | 0.74 |
| | +/0 & 0/0 | 62/176 | 1.62 (0.46–5.72) | 0.45 | +/0 & 0/0 | 49/141 | 0.86 (0.41–1.79) | 0.69 |
| | Null vs present [δ] | | 1.40 (0.78–2.50) | | Null vs present [δ] | | 1.34 (0.69–2.61) | |
| Ever smokers | | | | | | | | |
| | +/+ | 8/36 | 1.00 | | +/+ | 28/94 | 1.00 | |
| | +/0 | 48/158 | 1.57 (0.66–3.76) | | +/0 | 68/204 | 1.17 (0.69–1.98) | 0.70[e] |
| | 0/0 | 69/199 | 1.91 (0.82–4.47) | 0.12[e] | 0/0 | 29/79 | 1.13 (0.60–2.11) | |
| | +/0 & 0/0 | 117/357 | 1.77 (0.77–4.05) | 0.18 | +/0 & 0/0 | 97/283 | 1.16 (0.70–1.92) | |
| | Null vs present [δ] | | 1.31 (0.87–1.99) | | Null vs present [δ] | | 1.00 (0.61–1.67) | |
| **Frequency of smoking*** | | | | | | | | |
| ≤20 cigarettes smoked per day | | | | | | | | |
| | +/+ | 3/27 | 1.00 [b] | | +/+ | 18/61 | 1.00 [b] | |
| | +/0 | 34/100 | 3.06 (0.83–11.28) [b,c] | | +/0 | 39/133 | 0.85 (0.44–1.64) [b] | |
| | 0/0 | 46/128 | 3.35 (0.94–11.86) [b,d] | 0.12[e] | 0/0 | 18/50 | 0.87 (0.39–1.93) [b] | 0.72[e] |
| | +/0 & 0/0 | 80/228 | 3.25 (0.93–11.34)[b] | 0.07 | +/0 & 0/0 | 57/183 | 0.86 (0.46–1.60) [b] | 0.63 |
| | Null vs present [δ] | | 1.31 (0.78–2.21) [b] | | Null vs present [δ] | | 0.97 (0.50–1.87) [b] | |
| >20 cigarettes smoked per day | | | | | | | | |
| | +/+ | 5/9 | 1.0 [b] | | +/+ | 6/33 | 1.00 [b] | |
| | +/0 | 14/58 | 0.74 (0.21–2.63) [b] | | +/0 | 29/71 | 2.41 (0.83–6.99) | |
| | 0/0 | 23/71 | 0.86 (0.25–3.01) [b] | 0.95 [e] | 0/0 | 11/29 | 1.96 (0.61–6.34) | 0.34 [e] |
| | | | p-interaction[f] | 0.22 | | | p-interaction[f] | 0.21 |
| | +/0 – 0/0 | 37/129 | 0.86 (0.25–3.01) [b] | 0.72 | +/0 – 0/0 | 40/80 | 2.25 (0.81–6.25) | 0.12 |

[a]Unless otherwise stated, ORs were adjusted for age (continuous) and number of cigarettes smoked per day (continuous)

| | GSTM1 | | | | GSTT1 | | | |
|---|---|---|---|---|---|---|---|---|
| | **Genotype**‡ | **Cases/Controls** | **OR (95% CI)**[a] | **p-value** | **Genotype**‡ | **Cases/Controls** | **OR (95% CI)**[a] | **p-value** |
| | | | p-interaction[g] | 0.07 | | | p-interaction[g] | 0.17 |
| | Null vs present[δ] | | 1.11 (0.54–2.26)[b] | | Null vs present[δ] | | 1.00 (0.45–2.26)[b] | |

‡(+/+): wild-type genotype; (0/+): hemizygous deletion; (0/0): homozygous deletion

δNull vs present: Null = (0/0) versus reference group (present) comprised of **wild-type** (+/+) plus **hemizygotes** (+/0) genotypes

*a*Unless otherwise stated, ORs were adjusted for age (continuous) and number of cigarettes smoked per day (continuous)

*b*ORs adjusted for age (continuous)

*Ever smokers only

*c*p-value: 0.09

*d*p-value: 0.06

*e*p-value for trend

*f*p-interaction for hemizygous deletion (+/**0**)or homozygous deletion (**0/0**) and smoking intensity

*g*p-interaction combined hemizygous deletion (+/**0**) plus homozygous deletion (**0/0**) and smoking intensity

[a] Unless otherwise stated, ORs were adjusted for age (continuous) and number of cigarettes smoked per day (continuous)

[*] Ever smokers only

[b] ORs adjusted for age (continuous)

[b] ORs adjusted for age (continuous)

[b] ORs adjusted for age (continuous)

[c] p-value: 0.09

[b] ORs adjusted for age (continuous)

[b] ORs adjusted for age (continuous)

[d] p-value: 0.06

[e] p-value for trend

[b] ORs adjusted for age (continuous)

[e] p-value for trend

[b] ORs adjusted for age (continuous)

[b] ORs adjusted for age (continuous)

[δ] Null vs present: Null = (0/0) versus reference group (present) comprised of **wild-type** (+/+) plus **hemizygotes** (+/0) genotypes

[b] ORs adjusted for age (continuous)

[δ] Null vs present: Null = (0/0) versus reference group (present) comprised of **wild-type** (+/+) plus **hemizygotes** (+/0) genotypes

[b] ORs adjusted for age (continuous)

[b] ORs adjusted for age (continuous)

[b] ORs adjusted for age (continuous)

[b] ORs adjusted for age (continuous)

[b] ORs adjusted for age (continuous)

[e] p-value for trend

[e] p-value for trend

[f] p-interaction for hemizygous deletion (+/**0**)or homozygous deletion (**0/0**) and smoking intensity

[f] p-interaction for hemizygous deletion (+/**0**)or homozygous deletion (**0/0**) and smoking intensity

[b] ORs adjusted for age (continuous)

[g] p-interaction combined hemizygous deletion (+/**0**) plus homozygous deletion (**0/0**) and smoking intensity

[g] p-interaction combined hemizygous deletion (+/**0**) plus homozygous deletion (**0/0**) and smoking intensity

[δ] Null vs present: Null = (0/0) versus reference group (present) comprised of **wild-type** (+/+) plus **hemizygotes** (+/0) genotypes

[b] ORs adjusted for age (continuous)

[δ] Null vs present: Null = (0/0) versus reference group (present) comprised of **wild-type** (+/+) plus **hemizygotes** (+/0) genotypes

[b] ORs adjusted for age (continuous)