Published in final edited form as:

Cancer Detect Prev. 2009 ; 32(0): 403–415. doi:10.1016/j.cdp.2009.02.004.

Haplotype-tagging single nucleotide polymorphisms in the *GSTP1* gene promoter and susceptibility to lung cancer[★]

Xiang-Lin Tan^{a,b,*}, Roxana Moslehi^c, WeiGuo Han^{a,b}, and Simon D. Spivack^{a,b,d,e,f}

^aLaboratory of Human Toxicology and Molecular Epidemiology, Wadsworth Center, New York State Department of Health, Albany, NY, USA

^bDivision of Pulmonary Medicine, Albert Einstein College of Medicine, Bronx, NY 10461, USA

^cDepartment of Epidemiology and Biostatistics, School of Public Health, and Center for Excellence in Cancer Genomics, State University of New York, Rensselaer, NY, USA

^dDepartment of Epidemiology, Albert Einstein College of Medicine, Bronx, NY, USA

^eDepartment of Genetics, Albert Einstein College of Medicine, Bronx, NY, USA

^fDepartment of Environmental Health Sciences, School of Public Health, University at Albany, Albany, NY, USA

Abstract

Background—Glutathione S-transferase (GST) P1 is a major phase II xenobiotic-metabolizing enzyme in the human lung. Our laboratory had previously identified nine single nucleotide polymorphisms (SNPs) in the *GSTP1* gene promoter, which were then grouped into three main haplotypes (Hap1, Hap2, and Hap3) based on statistical inference. Hap3 was found to display a high expression phenotype. The main objective of the current study was to test the association between *GSTP1* promoter haplotypes with the risk of lung cancer after determining the promoter haplotypes experimentally through cloning and sequencing.

Methods—We conducted a case–control analysis of 150 subjects with lung cancer and 329 controls with no personal history of the disease. The three statistically inferred *GSTP1* promoter haplotypes were confirmed experimentally through cloning and sequencing. Haplotype-tagging SNPs were selected and *GSTP1* haplotypes were tested for genetic association to lung cancer using unconditional logistic regression after adjusting for confounders. Statistical interaction between *GSTP1* promoter haplotypes with either cigarette smoking or dietary fruit and vegetable intake were tested using the likelihood ratio test.

Results—We did not find protective effects of Hap3 against lung cancer, despite an adequately powered design for this main effect. Homozygous variants of tagSNPs –1738 T > A and –354 G > T, which tag Hap2, showed an increased (but statistically non-significant) risk of lung cancer among all subjects as well as among individuals with low fruit and vegetable intake, compared to homozygous wildtypes for these SNPs. We did not find significant interactions between Hap2 and dietary intake of fruits and vegetables.

[★]*Source of support:* This work was supported by NIH-R21 CA 94714 (to S.D. Spivack), NIH-R01 CA 10618 (to S.D. Spivack) and Prevent Cancer Foundation (Research Fellowship, to X.-L. Tan).

© 2009 Elsevier Ltd. All rights reserved.

^{*}Corresponding author at: Division of Pulmonary Medicine, Price Center for Genetic and Translational Medicine, Albert Einstein College of Medicine, 1301 Morris Park Ave., Bronx, NY 10461, USA. Tel.: +1 718 678 1503; fax: +1 718 678 1019. xtan@aecom.yu.edu (X.-L. Tan).

Conflict of interest statement

None declared.

Conclusions—Our results do not support significant main and modifying effects for *GSTP1* promoter haplotypes on susceptibility to lung cancer in this population, but reinforce the protective effects of dietary intake of fruits and vegetables.

Keywords

Lung cancer; *GSTP1*; Promoter polymorphisms; Gene–environment interaction

1. Introduction

Glutathione S-transferases (GSTs) are important phase II detoxification enzymes that catalyze the conjugation of electrophilic compounds with reduced glutathione. *GSTP1*, a member of the GST super family, has been identified as the major metabolic GST enzyme in human lung [1–6], and is one of the main detoxifiers of the activated form of benzo[*a*]pyrene (BaP) and other polycyclic aromatic hydrocarbons (PAHs), as well as a detoxifier of nitroso-family compounds, major carcinogenic metabolites derived from tobacco smoke [7–9]. Given that *GSTP1* is the most abundant GST isoform in the lung by RNA, protein and activity analyses [1–6], it is anticipated to be of particular importance in the detoxification of inhaled carcinogens. *GSTP1* polymorphisms could, therefore, be relevant as modifiers of lung cancer susceptibility, through their effect on the detoxification of epoxides and other reactive intermediates.

The *GSTP1* gene, mapped to chromosome 11q13 [10,11], is encoded in an allele-polymorphic gene locus, located in a 2.8-kb region, and consists of seven exons and six introns [3,12]. Several single nucleotide polymorphisms (SNPs) in the coding region of *GSTP1* have been identified. Two of the most studied polymorphisms, in exons 5 and 6, give rise to Ile105Val and Ala114Val amino acid substitutions and lead to diminished enzymatic activity [12–14]. In addition, the substitution in codon 105 has been shown to be associated with higher hydrophobic adduct levels in lung tissue and higher levels of PAH–DNA adducts in human lymphocytes [15]. A number of epidemiologic studies have explored possible associations between these two *GSTP1* polymorphisms and the risk of lung carcinoma, with conflicting results [16–21]. Recently, a meta-analysis of 130 studies showed that the Ile105Val and Ala114Val polymorphisms in the coding region of *GSTP1* gene do not have significant effects on the risk of lung cancer [22].

Our interest lies in the variants of the *GSTP1* gene promoter. Promoters are involved in the initiation of transcription and are therefore among the many important *cis*-acting elements that regulate gene expression. They may harbor polymorphisms that are potentially functionally relevant. One study investigated the association between the polymorphic –410 ATAAA repeat in the *GSTP1* promoter, and prostate cancer risk, but found no significant associations [23].

Our laboratory recently sequenced approximately 1.6 kb upstream of the distal and proximal promoter region of *GSTP1* translation initiation site, and identified 10 promoter polymorphisms: –1738 T > A, –1387 A > G, –830 G > C, –829 C > T, –566 G > C, –410 (ATAAA)_n, –354 G > T, 3289 T > Del, 3287 C > A, 3282 C > T. These SNPs were statistically grouped into three frequently observed haplotypes: Hap1 (43%), Hap2 (36%), and Hap3 (8%) [24]. Our group explored the interactions among these three main haplotypes containing the variations in the *GSTP1* gene promoter and candidate chemopreventive exposure in regulating human *GSTP1* expression; Hap3 displayed a high expression phenotype in reporter constructs, in peripheral lymphocytes, and in matched lung tissue, at the mRNA and protein level [24]. The association between these variants in the *GSTP1* promoter region and susceptibility to lung cancer has not yet been explored.

In this study, we aimed to confirm the three haplotypes in the *GSTP1* gene promoter by cloning and sequencing, and evaluate the association between the variants in *GSTP1* promoter region and lung cancer susceptibility, using the haplotype-tagging SNP approach, in a lung cancer case–control study. Given that high consumption of fruits and vegetables is associated with reduced risk of tobacco-induced lung cancer [25–33], we further tested for statistical interaction between these genetic variants and dietary nutrient intake, and/or tobacco smoking.

2. Materials and methods

2.1. Study subjects and data collection

We recruited 150 patients with a first diagnosis of lung cancer on the basis of pathologic examination from 2001 to 2006 at a tertiary care center (Albany Medical Center, AMC) and a large community hospital (St. Peter's Hospital) in Albany, NY. There were no recruitment restrictions on age, sex, race, or cancer stage. Controls were 329 cancer-free individuals recruited from three sources: (1) (Group 1) individuals undergoing lung sampling bronchoscopic procedures or surgery at AMC or St. Peter's Hospital ($n = 147$); (2) (Group 2) individuals seeking care in pulmonary, gastroenterology or cardiology clinics ($n = 87$); and (3) (Group 3) healthy controls employed at AMC hospital or New York State Department of Health employees ($n = 95$) living in the study region (Hudson Valley/Capitol region of New York State). Control subjects with any prior history of lung cancer were excluded from the study. This study was carried out under New York State Department of Health, AMC, and St. Peter's Hospital institutional review board-approved protocols. Written informed consent was obtained from each study participant.

At the first visit, which preceded any diagnostic procedures (bronchoscopy or surgery) and therefore preceded the diagnosis of lung cancer for cases and group 1 controls, all participants completed a research nurse-administered questionnaire, which elicits information on demographic factors, medical history, family history of lung or other malignancies, occupational exposure, and known or suspected risk factors for lung cancer. As part of the interview, participants were asked to provide information on frequency of fruit and vegetable consumption in the 12 months preceding the interview (cases and controls). The frequencies of fruit and vegetables for all participants are from 0 to 6 servings per day. For data analysis, the consumption of fruits and vegetables was categorized as low (0–2 servings/day), intermediate (3–4 servings/day) and high (5–6 servings/day). The fruit consumption category consisted of the intake of all varieties of fresh fruits, but excluded dried and canned fruits. The vegetable consumption category included cruciferous vegetables, such as broccoli, Brussel sprouts, cabbages, mustard greens, and cauliflower, as well as a query of others. Additionally, a lifetime smoking history for both active smoking and exposure to environmental tobacco smoke (ETS) was collected. An ever-smoker is someone who reported having smoked at least 100 cigarettes in their lifetime. Former smokers are subjects who quit smoking at least 3 months before the reference date, which was the date of interview for both cases and controls. Ever exposure to ETS in adulthood was defined as having been exposed for at least 1 h/day for at least 1 year due to the partner or at work. Pack-years of smoking (for active smoking) were calculated by dividing the average number of cigarettes smoked per day by 20 and multiplying the latter by the number of years of smoking for each phase. We further categorized the subjects as never smokers (0 pack-years), light smokers (≤ 10 pack-years) and heavy smokers (>10 pack-years) according to the pack-years of smoking. Occupational exposure was defined as either ever environmental exposure to significant sustained visible asbestos or silicate dusts, or sustained exposure to inhalable metals from grinding trades, for at least 1 year.

About 50% of the expected total number of eligible incident cases diagnosed during the study period were recruited. Most missed subjects were due to incomplete recruitment by clinicians in the hospitals, whereas only a few cases approached for participation refused to participate; participation rate among cases approached was 80%. Of the 160 cases and 340 controls who participated in the face-to-face interview, 156 (97.5%) cases and 332 (97.6%) controls each provided a 15-ml blood sample. Isolation of peripheral blood mononuclear cell (PBMC) samples was performed in standard fashion using a Ficoll gradient technique. Samples were stored in -180°C liquid nitrogen.

2.2. Experimental confirmation of the haplotypes in the *GSTP1* promoter by cloning and sequencing

A 1616-bp fragment of the *GSTP1* promoter region from -1804 to -199 (containing all the known SNP variants previously identified) was amplified by PCR using an Extensor Hi-Fidelity PCR Enzyme Mix (Abgene, Rochester, NY). The following oligomers were purchased from Operon Biotechnologies Inc. (Huntsville, AL, USA), for use in the amplification of the 1.6-kb region: the forward primer was 5'-GGG CTG AAT TTC CCC CGATAAT-3' and the reverse primer was 5'-GGA AGC CTT TCC CTC TTT CC-3'. The PCR thermal profile entailed an initial denaturation at 94°C for 2 min, and 34 cycles of denaturation at 94°C for 10 s, annealing at 62°C (5 cycles), 60°C (5 cycles) and 54°C (24 cycles) for 30 s, extension at 68°C for 2 min, followed by a final extension at 68°C for 10 min. PCR products were resolved by gel electrophoresis and were purified with a gel extraction kit (Qiagen, Valencia, CA). The PCR products were then ligated into pCR8/GW/TOPO[®] TA cloning vector (Invitrogen, Carlsbad, CA) and transfected into One Shot[®] Mach1[™]-T1R Chemically Competent *Escherichia coli* (Invitrogen). DNA from individual clones was sequenced on a PerkinElmer Biosystems ABI PRISM 377XL automated sequencer using the primers 5'-GTT GCA ACA AAT TGA TGA GCA ATG C-3' and 5'-GTT GCA ACA AAT TGA TGA GCA ATT A-3' provided by Invitrogen, and 5'-GGA ACA GAA AAC CAG CGA GAC-3' synthesized by Operon Biotechnologies, Inc. (Huntsville, AL, USA).

2.3. Selection of haplotype-based tagging SNPs

The principal hypothesis underlying this study was that there are one or more SNPs in the *GSTP1* promoter that are associated with an altered risk of lung cancer. We therefore aimed to identify a set of SNPs that efficiently tag all of the known common allelic variants [minor allele frequency (MAF) > 0.05] and are thus likely to tag most of the unknown common haplotypes in *GSTP1* promoter. We had initially identified the SNPs by sequencing approximately 1.6-kb upstream of the distal and proximal promoter region of *GSTP1* translation initiation site, using samples from 40 individuals [24]. For selection of tagSNPs, we incorporated the genotyping data from these 40 individuals into the Haploview program (<http://www.hapmap.org/haploview>). We identified a set of tagging SNPs of the *GSTP1* promoter with the following criteria: a minimal set of haplotypes that ensure an estimated r^2 of at least 0.8 to cover all possible haplotypes that had a frequency of at least 5% as evaluated by the programme Tagger function within Haploview (<http://www.broad.mit.edu/mpg/tagger/>). Four tagSNPs $-1738\text{ T} > \text{A}$, $-830\text{ G} > \text{C}$, $-566\text{ G} > \text{C}$, $-354\text{ G} > \text{T}$ were chosen to capture the common variants within this *GSTP1* promoter, with a minimum r^2 of 0.80 (Fig. 1).

2.4. Genotyping

Genomic DNA was isolated from PBMC samples using a standard, kit-based method (Gentra Systems, Minneapolis, MN) according to the manufacturer's recommendations. The DNA concentration was adjusted to $50\text{ }\mu\text{M}$ by TE buffer, and all DNA preparations were stored at 4°C until used for genotyping.

A TaqMan[®] (ABI) genotyping assay was employed to genotype all samples for three selected tagSNPs (–1738 T > A, –830 G > C, and –354 G > T). We performed genotyping for –830,829 GC > CT because of the complete linkage between –830 G > C and –829 C > T. For each of the SNPs, primer-probe sets (Table 1) were made using the Applied Biosystems design service (Foster City, CA). Two fluorogenic minor groove binder probes were used for each locus using the dyes 6-carboxyfluorescein (FAM; excitation, 494 nm) and VIC (excitation, 538 nm) which are easily differentiated in the Applied Biosystems Prism 7500HT PCR system. We carried out PCR on 10 ng genomic DNA using TaqMan[®] universal PCR master mix (Applied Biosystems, Foster City, CA), forward and reverse primers and FAM- and VIC-labeled probes. Real-time PCR was performed using 10.0 µl of TaqMan[®] 2× universal master mix (Applied Biosystems, Foster City, CA), 0.25 µl of primer-probe, 8.75 µl of RNase- and DNase-free water (Sigma), and 1 µl of sample DNA (50 µM). All assays were carried out in a 20-µl reaction in 96-well plates. Assay conditions were 10 min at 95 °C, and 40 cycles of 92 °C for 15 s and 58 °C for 1 min. Initially, the SNP assay was set up using SDS, Version 2.1, software (Applied Biosystems, Foster City, CA) as an absolute quantification assay, but after assay completion the plate was read using the allelic discrimination settings. Post-assay analysis was performed using the SDS software.

Because of the failure of TaqMan[®] (ABI) genotyping assay for the TagSNP –566 G > C, this SNP was determined by rapid capillary PCR followed by melting curve analysis using fluorescence-labeled hybridization probes in a Light-Cycler[®] thermocycler (Roche Diagnostics, Mannheim, Germany). The melting point analysis uses fluorescence resonance energy transfer to detect a polymorphic site [34]. The PCR primers and probes for melting point analysis are given in Table 1. Probes were designed and prepared by TIB MOLBIOL (Berlin, Germany). Analysis was performed in 10 µl volumes in glass capillaries (Roche Diagnostics) using Qiagen reagents: 1× PCR buffer, 2.5 mM total MgCl₂, 0.2 mM dNTPs, 0.5 U Hot Start Taq DNA polymerase, 0.5 µM of each primer, 0.15 µM of each probe, and 10 ng of DNA. Reaction conditions were as follows: initial denaturation at 95 °C for 10 min, then 45 cycles of denaturation at 95 °C for 15 s, annealing at 50 °C for 10 s and elongation 72 °C for 16 s with a ramping rate 0.4 °C/s. Melting curve analysis was performed with an initial denaturing step at 95 °C for 30 s, at 45 °C for 20 s slow heating to 75 °C with a ramping rate 0.1 °C/s and continuous fluorescence detection. The melting curve data were converted to melting peaks by plotting of the negative derivatives of fluorescence against temperature ($-dF/dT$).

For the purpose of quality control, a negative control containing all reagents, but with water instead of the DNA template was included in each amplification set. Genotyping was carried out blinded to case–control status. A 10% random sample was repeated to verify genotyping results, and 100% concordance was found.

2.5. Statistical analysis

Genotyping data were available from 150 cases and 329 controls. Departure of genotype frequencies from those expected under Hardy–Weinberg equilibrium was assessed among Caucasian controls by the asymptotic Pearson's Chi-square test with one degree of freedom. The Haploview (<http://www.hapmap.org/haploview>) program was used to assess the degree of linkage disequilibrium (LD) among SNPs according to the value of D' . Haplotypes of 150 cases and 329 controls were determined using the PHASE Version 2 software [35].

Chi-squared test were also used to compare the distribution of variables between case and controls. For analyses comparing cases and controls, odds ratios (ORs) and 95% confidence intervals (CIs) for association between each genotypes/haplotypes and the risk of lung cancer were obtained using unconditional logistic regression adjusting for gender, ethnicity, age at diagnosis, occupation risk, and pre-existing lung disease including emphysema,

COPD, chronic bronchitis, lung fibrosis, but not asthma. For main effect and interaction analyses of smoking, adjustments were made for dietary intake of fruits and vegetables as a continuous variable; while for main effect and interaction analyses of diet, pack-years of smoking (as a continuous variable) was also included in the logistic regression model. Trend statistics for risk of lung cancer in association with the genotypes were obtained using multivariate models. *P*-Values for tests of multiplicative interactions of genetic variants with the pack-years of smoking and dietary intake of fruits and vegetables were obtained by comparing logistic regression models with and without interaction term using likelihood-ratio tests. All tests were two-sided, and all statistical analyses were done by using the Statistical Analysis System software (Version 9.1; SAS Institute, Cary, NC).

3. Results

The sequence results for the three cloned regions representing the three main haplotypes (Hap1, Hap2 and Hap3) are illustrated in Fig. 2. Because of the –410 (ATAAA) pentanucleotide repeat variation, the downstream SNP sites, –354 G > T, –289 T > Del, –287 C > A, and –282 C > T, were sequenced using reverse primers. The sequencing results of cloned promoters provided experimental confirmation of the haplotypes, previously determined through statistical approaches, in the *GSTPI* promoter, as well as the homozygous and heterozygous states.

All SNPs evaluated appear to be in Hardy–Weinberg equilibrium among Caucasian controls. Frequencies of the variant alleles among Caucasian controls were 0.39 for *GSTPI* –1738A, 0.08 for –830,829CT (Hap3-defining), 0.48 for –566C, and 0.39 for –354T. The four investigated tagging SNPs were in LD in our study population ($r^2 = 0.89$). Haplotypes for these four *GSTPI* promoter SNPs (–1738 T > A, –830,829 GC > CT, –566 G > C, and –354 G > T) were statistically determined. TGCCG (Hap1), AGCGT (Hap2) and TCTGG (Hap3) haplotypes were the three most prevalent haplotypes, with haplotype frequencies of 0.48, 0.40 and 0.10, respectively, which is in concordance with our previous results drawn from the same population [24].

The characteristics of our study subjects were shown in Table 2. The mean age at diagnosis was 63.8 years for cases, 54.6 years for Group 1 controls and 51.2 years for all controls. Therefore, case were significantly older compared to all controls ($P < 0.01$) and compared to group 1 controls ($P < 0.01$). Cases were more likely to be male (54.7%), while there were more females among all controls (61.7%) and among group 1 controls (55.8%). Cases had higher proportion of smoking and low fruit and vegetable consumption, compared to all groups of controls ($P < 0.001$). A significantly higher proportion of cases (16.0%) than controls (8.5%) had occupational risk ($P = 0.02$); similarly, significantly higher proportion of cases (32.9%) than controls (17.5%) had pre-existing lung disease ($P < 0.001$). There were no significant differences in the distribution of pre-existing lung disease and occupation risk between group 1 controls and cases. Although the distribution of ethnicity was not significantly different between cases and controls, ethnicity was included in the logistic models because it is generally associated with the distribution of genotypes. Because the distribution of family history of lung cancer did not differ significantly between cases and controls, it was not included in the logistic models.

As expected, smoking and low consumption of fruits and vegetables were significant risk factors for lung cancer (Table 3). There was a strong dose–response relationship between lung cancer risk and the pack-years of smoking (P value for trend < 0.001) (Table 3). After adjustments for age at diagnosis, gender, ethnicity, pre-existing lung disease (exclusive of asthma) and consumption of fruit and vegetables (continuous), heavy smokers (> 10 pack-years) had a statistically significantly increased lung cancer risk (OR = 14.57, 95% CI 5.93–

43.01, $P < 0.001$), compared to never smokers. Conversely, high fruit and vegetable consumption was associated with a statistically significant decreased lung cancer risk (OR = 0.36, 95% CI 0.17–0.77, $P = 0.008$), compared with low fruit and vegetable intake. Comparison of cases and group 1 controls alone yielded similar results (ORs not shown in Table 3).

The association between each of the tagSNPs and haplotypes for the *GSTP1* gene promoter and the risk of lung cancer is shown in Table 4. In the multivariate analysis after adjustments for age at diagnosis, gender, ethnicity, pack-years of smoking, consumption of fruit and vegetable, and pre-existing lung disease, neither the –830, –829 GC > CT nor the –566 G > C SNP was associated with the risk of lung cancer. Homozygous variants at –1738 T > A had an increased risk of lung cancer compared to homozygous wildtypes (OR = 2.15, 95% CI 1.02–4.57, $P = 0.05$), although the difference did not reach statistical significance. Homozygous variants at –354 G > T also had an increased risk of lung cancer compared to homozygotes for the common alleles (OR = 2.10, 95% CI 0.99–4.43, $P = 0.05$), statistical significance was not reached (Table 4). Haplotypes analysis revealed a borderline increased risk of lung cancer for Hap2 homozygotes compared to individuals without Hap2 (OR = 2.15, 95% CI 1.02–4.53, $P = 0.05$) (Table 4).

Joint effects of *GSTP1* genotypes/haplotypes and dietary intake of fruits and vegetables on the risk of lung cancer are shown in Table 5. No statistically significant interaction was observed between the *GSTP1* genetic variants and dietary intake of fruits and vegetables in relation to lung cancer risk. Among subjects with low fruit and vegetable intake, significantly increased risk of lung cancer was observed among homozygous variants at –1738 T > A (OR = 2.98, 95% CI 1.16–7.68, $P = 0.02$) and –354 G > T (OR = 2.87, 95% CI 1.12–7.38, $P = 0.03$). Furthermore, Hap2 homozygotes had an increased risk of lung cancer among individuals with low fruit and vegetable intake (OR = 2.92, 95% CI 1.14–7.50, $P = 0.03$), but not among those who had high/medium fruit and vegetable intake (OR = 1.24, 95% CI 0.33–4.68, $P > 0.05$). Homozygous variants at –1738 T > A and –354 G > T occur in Hap2; therefore results of haplotype analysis are consistent with individual analyses of SNPs –1738 T > A and –354 G > T.

Joint effects of *GSTP1* genotypes/haplotypes and pack-years of smoking on the risk of lung cancer are shown in Table 6. Based on small numbers of case subjects, carriers of the –830,829CT variant allele who were never/light smokers had an increased risk of lung cancer (OR = 6.90, 95% CI 1.17–40.63, $P = 0.03$), compared to the referent category of never/light smokers who were homozygous wildtypes for –830,829 GC/GC. Accordingly, individuals with one or more copies of Hap3 seem to be at an increased risk of lung cancer, compared to the referent category (OR = 6.70, 95% CI 1.17–40.63, $P = 0.03$). Significant interaction was observed between –830,829 GC > CT SNPs and pack-years of smoking (P values for interaction = 0.04), as well as between Hap3 and pack-years of smoking, in relation to lung cancer risk (P values for interaction = 0.02) (Table 6).

4. Discussion

In this study, after experimentally confirming the three main haplotypes in the *GSTP1* promoter, we assessed the effects of the *GSTP1* gene promoter SNPs on individual susceptibility to lung cancer, using a haplotype-tagging SNP approach. We did not find a protective effect for Hap3, despite the previously published functional evaluation [24]. Homozygous variants at –1738 T > A or –354 G > T had a non-significantly increased risk of lung cancer, compared to homozygous wildtypes, and this elevated risk was more apparent in subjects with low fruit and vegetable intake. No statistically significant interaction was, however, observed between these tagging SNPs and dietary intake of fruit and vegetable.

A protective effect of high consumption of fruit and vegetable, particularly cruciferous vegetables, has been suggested for several cancers, including lung cancer [25–29]. The evidence for an inverse association between fruit and vegetable consumption and risk of lung cancer was also found in the present study. The biological rationale for this effect is related to the known role of plant-derived micronutrients and antioxidants in blocking cell damage due to oxidants [33,36]. Possible joint effects of fruit and vegetable consumption and genetic variants of phase II metabolism genes on lung cancer risk have been previously explored [30–32], with some positive results. In particular, London et al. [30] noted in a population of Chinese men that detectable levels of urinary isothiocyanate (ITC), corresponding to high intakes of dietary ITCs (e.g., in crucifer vegetables), reduced lung cancer risk only among individuals with the *GSTM1* and/or *GSTT1* null genotype, where clearance of the ITCs is impaired and half-life of these protective factors prolonged. Recently, Sørensen et al. [32], evaluated interactions between *GSTM1*, *GSTT1*, and *GSTP1* polymorphisms, smoking and intake of fruits and vegetables in relation to lung cancer. That group found that fruit and vegetable consumption reduced the risk of lung cancer only among carriers of at least one functional *GSTM1* allele, and among carriers of two *GSTP1* Val alleles. Clearly, the kinetics is complex, in that presumptive inducers of phase II metabolism such as crucifers are themselves cleared by that metabolism.

Our previous study showed that Hap2-related *GSTP1* gene expression is low in both blood and lung tissue, compared to Hap3 [24]. Although the association between *GSTP1* gene expression and lung cancer risk is still not well understood, there might be an accumulation of phase I-activated smoking carcinogens by slowing the quenching or clearance because of the lower *GSTP1* (phase II) expression, thus increasing the risk of lung cancer. Ritchie et al. recently reported that *GSTP*-null mice (i.e. those with lower levels of GST activity) were found to have substantially increased numbers of lung adenomas relative to wildtype mice following exposure to BaP, 3-methylcholanthrene, and urethane, suggesting that variability in *GSTP* expression or function influences lung tumorigenesis [37].

In addition, it is possible that the promoter SNPs –1738 T > A, –354 G > T simply are indicators of, or linked to, more important SNPs in the coding region, which are mechanistically involved in an altered enzyme function [10–12]. This may be supported by the fact that Hap1 individuals, previously shown to have mRNA and protein levels equivalent to those of Hap2, had no higher lung cancer risk. Several future approaches to address this question are worth considering. First, systematic structure–activity relationship studies could be conducted on the effects of these SNPs on promoter function in the *GSTP1* gene. Second, volunteer studies could be conducted in subjects who have the same genotype in the coding region but who differ as to promoter genotype. A third possibility would be to study patient groups that are sufficiently large to allow detection of additional effects of other upstream promoter SNPs beyond those of the current 1.6 kb promoter region SNPs. A definitive answer is likely to require a combination of all three approaches.

No evidence was found for an association between tagSNPs –830,829 GC > CT (defining Hap3) and –566 C > T (Hap1) with the risk of lung cancer among all subjects, prior to stratification. The study had over 80% power to detect a 2.0-fold increase in risk of lung cancer for a dominant allele with allele frequency of 0.10 at the 5% significance level. We found that uniquely among subjects who were never/light smokers, tagSNPs –830,829 GC > CT and Hap3 had an increased risk of lung cancer; this association was not observed among heavy smokers (*P* value for interaction = 0.04). However, because of wide confidence intervals, our results should be interpreted with caution.

This unexpected finding that there may be a possible increased lung cancer risk with Hap3 copy number exclusive to non-smokers may indeed be a chance association, as lung cancer

mechanisms in the predominantly female never-smoking subgroup may entail different mechanisms than that of the smokers. To our knowledge, the effect of interaction between SNPs in the *GSTP1* gene promoter and smoking related on the risk of lung cancer has not been previously reported. Sørensen et al. [32] found no significant interactions between SNPs in the *GSTP1* coding region and smoking in relation to lung cancer. They offered the plausible explanation that the strong association of smoking with lung cancer may override modest effects of genotypes on the development of lung cancer [38,39].

Some potential power limitations and sources of bias of the present study should be considered. Despite having adequate power to exclude main effects of OR > 2.0, we cannot exclude a smaller effect or chance findings for borderline associations. We tested for statistical interaction between *GSTP1* promoter SNPs, diet and tobacco smoking. However, we could not perform a three-way interaction test for tobacco smoke × diet × *GSTP1* promoter SNPs because of the sample size restrictions. Although our cases and controls were not matched on age and gender; we included these factors as confounding factors in our analysis. We conducted our case–control comparisons by considering the three groups of controls separately to account for differences in age and variables; the results of individual control group analyses were similar to the analyses with all controls combined. We also adjusted for other potential lung cancer risk factors including ethnicity, occupation risk, pre-existing lung disease and consumption of fruit and vegetable, pack-years of smoking in our analyses.

While recall bias is a potential problem in most case–control studies, it is highly unlikely to affect the results of our study, as the patient recruitment preceded the operation (surgery or bronchoscopy) that subsequently yielded the pathologic diagnosis by several days to 2 weeks. Therefore, recall occurred before diagnosis among all cases, and the majority of the controls (Group 1). Further, when we compared case–control analyses entailing all controls (those undergoing, as well as those not undergoing, biopsy procedures who knew their control status, as did the interviewing nurse), *versus* only those controls undergoing biopsy procedures (who were thus “time-blinded” to their diagnosis, as was the nurse), the results were virtually comparable; although, the analysis was underpowered. Therefore, recall or actual modification of dietary habits after a cancer diagnosis (or non-diagnosis, for controls) is not likely to be relevant. Additionally, a detailed sensitivity analysis including all of these cases and controls indicated that potential selection bias is unlikely to have a major impact on the patterns observed in our study.

In summary, despite previous *in vitro* and *in vivo* functional data suggestive of Hap3 having a high expression phenotype, we found no definite evidence for protection from lung cancer. Given the availability of detailed knowledge of both the *GSTP1* promoter haplotype structure, and the *in vitro* and *in vivo* functional effects of these variants, further epidemiologic studies with a larger population, as well as more extensive functional studies on larger regions of the promoter contiguous with coding sequence, may further clarify whether and how *GSTP1* promoter polymorphisms influence the risk of lung cancer associated with diet or smoking.

Acknowledgments

We wish to thank all participants, the research nurse staff of AMC (Angela Sheehan, RN, Kathy Mokhiber, RN, Ann Venezia, RN, Paula Malone, RN) for their contribution to the data collection, the participating clinical colleagues in Pulmonary & Critical Care Medicine and Thoracic Surgery at AMC, and Ms. Adrienne Hoyt and Dr. Ranjana Singh at the School of Public Health, SUNY-Albany.

References

1. Awasthi YC, Singh SV, Ahmad H, Moller PC. Immunocytochemical evidence for the expression of *GST1*, *GST2*, and *GST3* gene loci for glutathione S-transferase in human lung. *Lung*. 1987; 165(6): 323–32. [PubMed: 3123804]
2. Di Ilio C, Del Boccio G, Aceto A, Casaccia R, Mucilli F, Federici G. Elevation of glutathione transferase activity in human lung tumor. *Carcinogenesis*. 1988; 9(2):335–40. [PubMed: 3123087]
3. Moscow JA, Fairchild CR, Madden MJ, Ransom DT, Wieand HS, O'Brien EE, et al. Expression of anionic glutathione-S-transferase and P-glycoprotein genes in human tissues and tumors. *Cancer Res*. 1989; 49(6):1422–8. [PubMed: 2466554]
4. Anttila S, Hirvonen A, Vainio H, Husgafvel-Pursiainen K, Hayes JD, Ketterer B. Immunohistochemical localization of glutathione S-transferases in human lung. *Cancer Res*. 1993; 53(23):5643–8. [PubMed: 8242618]
5. Coles B, Yang M, Lang NP, Kadlubar FF. Expression of *hGSTP1* alleles in human lung and catalytic activity of the native protein variants towards 1-chloro-2,4-dinitrobenzene, 4-vinylpyridine and (+)-anti benzo[a]pyrene-7,8-diol-9,10-oxide. *Cancer Lett*. 2000; 156(2):167–75. [PubMed: 10880766]
6. Spivack SD, Hurteau GJ, Fasco MJ, Kaminsky LS. Phase I and II carcinogen metabolism gene expression in human lung tissue and tumors. *Clin Cancer Res*. 2003; 9(16 Pt 1):6002–11. [PubMed: 14676126]
7. Nakajima T, Elovaara E, Anttila S, Hirvonen A, Camus AM, Hayes JD, et al. Expression and polymorphism of glutathione S-transferase in human lungs: risk factors in smoking-related lung cancer. *Carcinogenesis*. 1995; 16(4):707–11. [PubMed: 7728947]
8. Fields WR, Morrow CS, Doss AJ, Sundberg K, Jernström B, Townsend AJ. Overexpression of stably transfected human glutathione S-transferase P1-1 protects against DNA damage by benzo[a]pyrene diol epoxide in human T47D cells. *Mol Pharmacol*. 1998; 54(2):298–304. [PubMed: 9687571]
9. Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst*. 1999; 91(14):1194–210. [PubMed: 10413421]
10. Suzuki T, Coggan M, Shaw DC, Board PG. Electrophoretic and immunological analysis of human glutathione S-transferase isozymes. *Ann Hum Genet*. 1987; 51(Pt 2):95–106. [PubMed: 3118757]
11. Suzuki T, Johnston PN, Board PG. Structure and organization of the human alpha class glutathione S-transferase genes and related pseudogenes. *Genomics*. 1993; 18(3):680–6. [PubMed: 8307579]
12. Ali-Osman F, Akande O, Antoun G, Mao JX, Buolamwini J. Molecular cloning, characterization, and expression in *Escherichia coli* of full-length cDNAs of three human glutathione S-transferase Pi gene variants. Evidence for differential catalytic activity of the encoded proteins. *J Biol Chem*. 1997; 272(15):10004–12. [PubMed: 9092542]
13. Zimniak P, Nanduri B, Pikula S, Bendorowicz-Pikuła J, Singhal SS, Srivastava SK, et al. Naturally occurring human glutathione S-transferase GSTP1-1 isoforms with isoleucine and valine in position 104 differ in enzymic properties. *Eur J Biochem*. 1994; 224(3):893–9. [PubMed: 7925413]
14. Watson MA, Stewart RK, Smith GB, Massey TE, Bell DA. Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. *Carcinogenesis*. 1998; 19(2):275–80. [PubMed: 9498276]
15. Butkiewicz D, Grzybowska E, Phillips DH, Hemminki K, Chorazy M. Polymorphisms of the *GSTP1* and *GSTM1* genes and PAH-DNA adducts in human mononuclear white blood cells. *Environ Mol Mutagen*. 2000; 35(2):99–105. [PubMed: 10712743]
16. Ryberg D, Skaug V, Hewer A, Phillips DH, Harries LW, Wolf CR, et al. Genotypes of glutathione transferase M1 and P1 and their significance for lung DNA adduct levels and cancer risk. *Carcinogenesis*. 1997; 18(7):1285–9. [PubMed: 9230269]
17. Jourenkova-Mironova N, Wikman H, Bouchardy C, Voho A, Dayer P, Benhamou S, et al. Role of glutathione S-transferase *GSTM1*, *GSTM3*, *GSTP1* and *GSTT1* genotypes in modulating susceptibility to smoking-related lung cancer. *Pharmacogenetics*. 1998; 8(6):495–502. [PubMed: 9918133]

18. Harris MJ, Coggan M, Langton L, Wilson SR, Board PG. Polymorphism of the Pi class glutathione S-transferase in normal populations and cancer patients. *Pharmacogenetics*. 1998; 8(1):27–31. [PubMed: 9511178]
19. To-Figueras J, Gené M, Gómez-Catalán J, Piqué E, Borrego N, Carrasco JL, et al. Genetic polymorphism of glutathione S-transferase P1 gene and lung cancer risk. *Cancer Causes Control*. 1999; 10(1):65–70. [PubMed: 10334644]
20. Miller DP, Neuberg D, de Vivo I, Wain JC, Lynch TJ, Su L, et al. Smoking and the risk of lung cancer: susceptibility with GSTP1 polymorphisms. *Epidemiology*. 2003; 14(5):545–51. [PubMed: 14501269]
21. Miller DP, De Vivo I, Neuberg D, Wain JC, Lynch TJ, Su L, et al. Association between self-reported environmental tobacco smoke exposure and lung cancer: modification by *GSTP1* polymorphism. *Int J Cancer*. 2003; 104(6):758–63. [PubMed: 12640684]
22. 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(4):e91. [PubMed: 16509765]
23. Platz EA, Krithivas K, Kantoff PW, Stampfer MJ, Giovannucci E. ATAAA repeat upstream of glutathione S-transferase P1 and prostate cancer risk. *Urology*. 2002; 59(1):159–64. [PubMed: 11796315]
24. Cauchi S, Han W, Kumar SV, Spivack SD. Haplotype–environment interactions that regulate the human glutathione S-transferase P1 promoter. *Cancer Res*. 2006; 66(12):6439–48. [PubMed: 16778223]
25. Goldbohm RA, Voorrips LE. Epidemiology of nutrition and lung cancer. *Nestle Nutr Workshop Ser Clin Perform Program*. 2000; 4:23–35.
26. Brennan P, Butler J, Agudo A, Benhamou S, Darby S, Fortes C, et al. Joint effect of diet and environmental tobacco smoke on risk of lung cancer among nonsmokers. *J Natl Cancer Inst*. 2000; 92(5):426–7. [PubMed: 10699073]
27. Darby S, Whitley E, Doll R, Key T, Silcocks P. Diet, smoking and lung cancer: a case–control study of 1000 cases and 1500 controls in SouthWest England. *Br J Cancer*. 2001; 84(5):728–35. [PubMed: 11237398]
28. Männistö S, Smith-Warner SA, Spiegelman D, Albanes D, Anderson K, van den Brandt PA, et al. Dietary carotenoids and risk of lung cancer in a pooled analysis of seven cohort studies. *Cancer Epidemiol Biomark Prev*. 2004; 13(1):40–8.
29. Rylander R, Axelsson G. Lung cancer risks in relation to vegetable and fruit consumption and smoking. *Int J Cancer*. 2006; 118(3):739–43. [PubMed: 16108070]
30. 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(9231):724–9. [PubMed: 11085692]
31. Wang LI, Giovannucci EL, Hunter D, Neuberg D, Su L, Christiani DC. Dietary intake of cruciferous vegetables. Glutathione S-transferase (GST) polymorphisms and lung cancer risk in a Caucasian population. *Cancer Causes Control*. 2004; 15(10):977–85. [PubMed: 15801482]
32. Sørensen M, Raaschou-Nielsen O, Brasch-Andersen C, Tjønneland 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(2):137–44. [PubMed: 17123660]
33. Loft S, Poulsen HE. Cancer risk and oxidative DNA damage in man. *J Mol Med*. 1996; 74(6):297–312. [PubMed: 8862511]
34. Chang-Claude J, Popanda O, Tan XL, Kropp S, Helmbold I, von Fournier D, et al. Association between polymorphisms in the DNA repair genes, *XRCC1*, *APE1*, and *XPB* and acute side effects of radiotherapy in breast cancer patients. *Clin Cancer Res*. 2005; 11(13):4802–9. [PubMed: 16000577]
35. Stephens M, Donnelly P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet*. 2003; 73(5):1162–9. [PubMed: 14574645]
36. Ritchie KJ, Henderson CJ, Wang XJ, Vassieva O, Carrie D, Farmer PB, et al. Glutathione transferase pi plays a critical role in the development of lung carcinogenesis following exposure to

- tobacco-related carcinogens and urethane. *Cancer Res.* 2007; 67(19):9248–57. [PubMed: 17909032]
37. Lampe JW. Health effects of vegetables and fruit: assessing mechanisms of action in human experimental studies. *Am J Clin Nutr.* 1999; 70(3 Suppl):475S–90S. [PubMed: 10479220]
38. Wakai K, Inoue M, Mizoue T, Tanaka K, Tsuji I, Nagata C, et al. Tobacco smoking and lung cancer risk: an evaluation based on a systematic review of epidemiological evidence among the Japanese population. *Jpn J Clin Oncol.* 2006; 36(5):309–24. [PubMed: 16735374]
39. Schwartz AG, Prysak GM, Bock CH, Cote ML. The molecular epidemiology of lung cancer. *Carcinogenesis.* 2007; 28(3):507–18. [PubMed: 17183062]

Fig. 1. Linkage disequilibrium map of *GSTP1* promoter single nucleotide polymorphisms and haplotype frequencies among initial 40 individuals in our study population. The red squares indicate a statistically significant allelic association between the paired of SNPs. Numbers inside the square are the value of D' multiplied by 100. $D' = 1$ indicates complete linkage disequilibrium, while $D' = 0$ indicates complete linkage equilibrium. Numbers on the right represent haplotype frequencies (e.g., 0.432 = 43.2%).

Fig. 2.
 The sequence chromatogram tracing results of the cloned *GSTP1* promoter haplotypes (Hap1, Hap2, and Hap3). (*) indicates that -354 G > T, -289 T > Del, -287 C > A and -282 C > T were sequenced using a reverse primer.

Table 1

Primers and probes used for genotyping of tagging single nucleotide polymorphisms.

Varinats	dbSNP #	Genotyping method	Primer or probe sequence (5'-3') ^a
-1738T>A	rs6591255	TaqMan genotyping assay	F: CTGAATTTCCCGGATAATAAGTCTCTTAAA R: CACCGAAGACACGAGTTTACTTATGT P1: VIC-CCCAAAACAGAAATAA-NFQ P2: FAM-CCCAAAACTGAAATAA-NFQ
-830,829 GC>CT	rs34743590	TaqMan genotyping assay	F: CGCACACACCTATGTAACAAACCT R: CTGGCCGTCCAACCTTTATTGAAG P1: VIC-CATCCTGCACATGTAC-NFQ P2: FAM-CACATCCTCTACATGTAC-NFQ
-354G>T	rs 17593068	TaqMan genotyping assay	F: CCTCTCCCTGCCCTGT R: CGCGCCGGGACATC P1: VIC-TTGCACACCGCTTC-NFQ P2: FAM-TTGCACAACCGCTTC-NFQ
-566G>C	rs694746	LightCycler genotyping assay	F: GATGTGGCACGCACCTATAA R: GACGTAGTCTCACGCTGTGG A: LC Red 640-CCACTGCACTCCAGCCTGGG-ph S: TGAGCCGCCAAGATCGC-FITC

^aF, forward primer; R, reverse primer; P1, probe 1; P2, probe 2; A, anchor; S, sensor; NFQ, nonfluorescent quencher. Polymorphic site is shaded.

Table 2

Characteristics of the study subjects.

	Cases (n = 150)		Controls (n = 329)		Chi-square P value
	n	%	n	%	
<i>Gender^a</i>					
Male	82	54.7	123	37.4	<0.01
Female	68	45.3	203	61.7	
<i>Ethnicity^b</i>					
Caucasian	126	84.0	281	85.4	0.20
Others	8	5.3	30	9.1	
<i>Age (years)^c</i>					
40	3	2.0	74	22.5	<0.001
41–50	10	6.7	77	23.4	
51–60	44	29.3	84	25.5	
61–70	47	31.3	47	14.3	
71	43	28.7	38	11.6	
<i>Smoking status</i>					
Never smokers	6	4.0	103	31.3	<0.001
Former smokers	96	64.0	132	40.1	
Current smokers	48	32.0	94	28.6	
<i>Smoke dose (lifetime pack-years)^d</i>					
Never active	5	3.3	103	31.3	<0.001
<10	8	5.3	65	19.8	
10–19	9	6.0	32	9.7	
20–29	11	7.3	35	10.6	
30+	116	77.3	92	28.0	
<i>Family history of lung cancer^e</i>					
Yes (first degree)	32	21.3	50	15.2	0.17
Yes (second degree)	9	6.0	30	9.1	
No	109	72.7	246	74.8	

		<u>Cases (n = 150)</u>		<u>Controls (n = 329)</u>		<i>Chi-square P value</i>
		n	%	n	%	
<i>Fruits and vegetable consumption^f</i>						
High		16	10.7	78	23.7	<0.001
Intermediate		23	15.3	87	26.4	
Low		105	70.0	148	45.0	
<i>Occupation risk^g</i>						
Yes		24	16.0	28	8.5	0.02
No		126	84.0	300	91.2	
<i>Pre-existing lung disease^h</i>						
No		100	67.1	245	82.5	<0.01
Yes		49	32.9	52	17.5	

^aData missing for 3 controls.

^bData missing for 16 cases and 18 controls.

^cData missing for 3 cases and 9 controls.

^dData missing for 1 case and 2 controls.

^eData missing for 3 controls.

^fData missing for 6 cases and 16 controls.

^gData missing for 1 case and 1 control.

^h1 case and 32 controls with asthma not included.

Table 3

Risk of lung cancer associated with smoking and with fruit and vegetable consumption.

Possible factors to lung cancer	Cases (n= 150) ^a		Controls (n= 329) ^a		OR (95% CI)	Adjusted OR (95% CI) ^b
	n	%	n	%		
<i>Smoking status</i>						
Never smokers	6	4.00	103	31.31	1.00	1.00
Former smokers	96	64.00	132	40.12	12.48 (5.26–29.62) ^c	7.56 (2.81–20.36) ^c
Current smokers	48	32.00	94	28.57	8.76 (3.59–21.42) ^c	8.80 (3.08–25.14) ^c
<i>P</i> _{trend}					<0.001	<0.001
<i>Smoking dose</i>						
Never smokers (0 pack-years)	5	3.36	103	31.50	1.00	1.00
Light smokers (10 pack-years)	9	6.04	76	23.24	2.07 (0.71–6.07)	2.30 (0.62–8.60)
Heavy smokers (>10 pack-years)	135	90.60	148	45.26	15.96 (6.79–37.53) ^c	14.57 (5.93–43.01) ^c
<i>P</i> _{trend}					<0.001	<0.001
<i>Fruit and vegetable consumption</i>						
Low (0–2 servings/day)	104	72.22	151	48.24	1.00	1.00
Medium (3–4 servings/day)	24	16.67	85	27.16	0.43 (0.25–0.72) ^d	0.66 (0.34–1.25)
High (5–6 servings/day)	16	11.11	77	24.60	0.32 (0.18–0.57) ^c	0.36 (0.17–0.77) ^d
<i>P</i> _{trend}					0.007	0.004

^aNumbers in individual cells may not add up to this total because of missing data.

^bAdjusted for age at diagnosis (continuous), gender (male and female), ethnicity (Caucasian and non-Caucasian), occupation risk (yes and no), pre-existing lung disease (yes and no; exclusive of asthma) and consumption of fruit and vegetable (continuous; for smoking status and smoking dose), pack-years of smoking (continuous; for fruits and vegetable consumption).

^c*P*value < 0.001.

^d*P*value < 0.01.

Table 4

Association between the tagging SNP and haplotype of the *GSTP1* gene promoter and the risk for lung cancer.

Genotype/haplotype	Cases		Controls		OR (95% CI)	Adjusted OR (95% CI) ^a
	n	%	n	%		
<i>-1738 T>A</i>						
TT	52	34.7	113	34.3	1.00	1.00
TA	64	42.7	175	53.2	0.80 (0.51–1.23)	0.79 (0.45–1.38)
AA	34	22.7	41	12.5	1.80 (1.02–3.16) ^b	2.15 (1.02–4.57) ^b
<i>P</i> _{trend}					0.14	0.15
<i>-830,829 GC>CT</i>						
GC/GC	123	82.0	269	81.8	1.00	1.00
GC/CT	27	18.0	60	18.2	0.99 (0.59–1.67)	0.86 (0.43–1.73)
CT/CT					0.88 (0.17–4.57)	1.04 (0.08–14.50)
<i>P</i> _{trend}					0.92	0.72
<i>-566 G>C</i>						
GG	46	30.7	81	24.6	1.00	1.00
GC	70	46.7	170	51.7	0.73 (0.46–1.15)	0.66 (0.36–1.23)
CC	34	22.7	78	23.7	0.77 (0.45–1.32)	0.78 (0.39–1.60)
<i>P</i> _{trend}					0.31	0.51
<i>-354 G>T</i>						
GG	53	35.3	114	34.7	1.00	1.00
GT	63	42.0	174	52.9	0.78 (0.50–1.20)	0.75 (0.43–1.31)
TT	34	22.7	41	12.5	1.78 (1.02–3.12) ^b	2.10 (0.99–4.43) ^b
<i>P</i> _{trend}					0.16	0.19
<i>TGCCG (Hap1)</i>						
0	47	31.3	81	24.6	1.00	1.00
1	70	46.7	174	52.9	0.69 (0.44–1.09)	0.64 (0.35–1.18)
2	33	22.0	74	22.5	0.77 (0.45–1.33)	0.78 (0.38–1.60)
<i>P</i> _{trend}					0.30	0.50
<i>AGCGT (Hap2)</i>						
0	53	35.3	118	35.9	1.00	1.00

Genotype/haplotype	Cases		Controls		OR (95% CI)	Adjusted OR (95% CI) ^a
	n	%	n	%		
1	63	42.0	170	51.7	0.83 (0.53–1.27)	0.78 (0.45–1.36)
2	34	22.7	41	12.5	1.85 (1.06–3.22) ^b	2.15 (1.02–4.53) ^b
P _{trend}					0.11	0.16
TCTGG (Hap3)						
0	125	83.3	269	81.8	1.00	1.00
1	23	15.3	56	17.0	0.88 (0.52–1.50)	0.78 (0.38–1.61)
2	2	1.3	4	1.2	1.08 (0.19–5.95)	1.10 (0.08–16.26)
P _{trend}					0.73	0.57

^aAdjusted for age at diagnosis (continuous), gender (male and female), ethnicity (Caucasian and non-Caucasian), pack-years of smoking (continuous), consumption of fruit and vegetable (continuous), occupation risk (yes and no), and pre-existing lung disease (yes and no; exclusive of asthma).

^bP_{value} 0.05.

Table 5

Joint effects of the dietary intake of fruits and vegetables and *GSTP1* genotypes/haplotypes on the risk of lung cancer.

Genotype/haplotype	Low fruit/vegetable intake		High/medium fruit/vegetable intake		<i>P</i> _{Interaction}
	No. ^a	Adjusted OR (95% CI) ^b	No. ^a	Adjusted OR (95% CI) ^b	
<i>-1738 T> A</i>					
TT	34/52	1.00	15/54	1.00	
TA	45/80	0.75 (0.37–1.54)	18/88	0.99 (0.37–2.63)	
AA	25/19	2.98 (1.16–7.68) ^c	7/20	1.21 (0.32–4.59)	0.37
<i>P</i> _{trend}		0.08		0.82	
<i>-830,829GC> CT</i>					
GC/GC	86/130	1.00	32/126	1.00	
GC/CT	16/18	1.08 (0.42–2.75)	8/34	0.72 (0.23–2.26)	
CT/CT	2/3	1.01 (0.07–15.28)	0/2	NA	0.49
<i>P</i> _{trend}		0.89		0.55	
<i>-830,829GC> CT</i>					
GC/GC	86/130	1.00	32/126	1.00	
GC and CT (grouped)	18/21	1.07 (0.44–2.62)	8/36	0.71 (0.23–2.24)	0.49
<i>-566 G> C</i>					
GG	33/35	1.00	11/43	1.00	
GC	49/78	0.59 (0.27–1.28)	18/82	0.92 (0.32–2.69)	
CC	22/38	0.60 (0.24–1.47)	11/37	1.12 (0.34–3.74)	0.30
<i>P</i> _{trend}		0.26		0.85	
<i>-354 G> T</i>					
GG	35/53	1.00	15/54	1.00	
GT	44/79	0.70 (0.34–1.42)	18/88	0.97 (0.37–2.59)	
TT	25/19	2.87 (1.12–7.38) ^c	7/20	1.20 (0.32–4.57)	0.22
<i>P</i> _{trend}		0.11		0.84	
<i>TGCCG (Hap I)</i>					
0	34/35	1.00	11/43	1.00	
1	48/79	0.56 (0.26–1.20)	19/85	0.92 (0.32–2.69)	
2	22/37	0.60 (0.24–1.47)	10/34	1.13 (0.33–3.84)	0.32

Genotype/haplotype	Low fruit/vegetable intake		High/medium fruit/vegetable intake		$P_{\text{interaction}}$
	No. ^a	Adjusted OR (95% CI) ^b	No. ^a	Adjusted OR (95% CI) ^b	
P_{trend}		0.25		0.85	
<i>AGCGT</i> (Hap2)					
0	35/54	1.00	15/57	1.00	
1	44/78	0.72 (0.35–1.46)	18/85	1.02 (0.39–2.71)	
2	25/19	2.92 (1.14–7.50) ^c	7/20	1.24 (0.33–4.68)	0.43
P_{trend}		0.10		0.79	
<i>TCTGG</i> (Hap3)					
0	87/130	1.00	33/126	1.00	
1	15/19	1.00 (0.38–2.59)	7/34	0.66 (0.20–2.14)	
2	2/2	1.10 (0.07–18.32)	0/2	NA	0.41
P_{trend}		0.97		0.47	
<i>TCTGG</i> (Hap3)					
0	87/130	1.00	32/126	1.00	
1 and 2 (grouped)	17/21	1.01 (0.41–2.51)	7/36	0.65 (0.20–2.11)	0.45

^aNumber of cases/controls.

^bAdjusted for age at diagnosis (continuous), gender (male and female), ethnicity (Caucasian and non-Caucasian), pack-years of smoking (continuous), occupation risk (yes and no), and pre-existing lung disease (yes and no; exclusive of asthma); NA = not applicable because of small numbers.

^c $P_{\text{value}} < 0.05$.

Table 6
Joint effects of pack-years of smoking and *GSTP1* genotypes/haplotypes on the risk of lung cancer.

Genotype/haplotype	Never/light smokers		Heavy smokers		<i>P</i> _{Interaction}
	No. ^a	Adjusted OR (95% CI) ^b	No. ^a	Adjusted OR (95% CI) ^b	
-1738 <i>T</i> > <i>A</i>					
TT	4/67	1.00	47/45	1.00	
TA	8/92	2.18 (0.35–13.51)	56/82	0.63 (0.34–1.18)	
AA	2/20	2.69 (0.21–33.89)	32/21	1.79 (0.76–4.23)	0.52
<i>P</i> _{trend}		0.37		0.46	
-830,829 <i>GC</i> > <i>CT</i>					
GC/GC	8/147	1.00	115/120	1.00	
GC/CT	6/27	7.07 (1.22–41.16) ^c	18/28	0.66 (0.30–1.45)	
CT/CT	0/5	NA	2/0	NA	0.04
<i>P</i> _{trend}		0.05		0.48	
-830,829 <i>GC</i> > <i>CT</i>					
GC/GC	8/147	1.00	115/120	1.00	
GC and CT (grouped)	6/32	6.90 (1.17–40.63) ^c	20/28	0.70 (0.32–1.53)	0.04
-566 <i>G</i> > <i>C</i>					
GG	6/42	1.00	39/39	1.00	
GC	6/92	0.26 (0.05–1.47)	64/77	0.81 (0.41–1.63)	
CC	2/45	0.18 (0.02–1.77)	32/32	1.07 (0.48–2.39)	0.10
<i>P</i> _{trend}		0.09		0.86	
-354 <i>G</i> > <i>T</i>					
GG	4/68	1.00	48/45	1.00	
GT	8/91	2.30 (0.37–14.16)	55/82	0.60 (0.32–1.12)	
TT	2/20	2.80 (0.22–35.24)	32/21	1.76 (0.75–4.13)	0.50
<i>P</i> _{trend}		0.35		0.52	
<i>TGCCG</i> (<i>Hap</i> <i>h</i>)					
0	6/42	1.00	40/39	1.00	
1	6/95	0.25 (0.04–1.43)	64/78	0.78 (0.39–1.55)	
2	2/42	0.18 (0.02–1.84)	31/31	1.03 (0.46–2.30)	0.11

Genotype/haplotype	Never/light smokers		Heavy smokers		$P_{\text{interaction}}$
	No. ^a	Adjusted OR (95% CI) ^b	No. ^a	Adjusted OR (95% CI) ^b	
P_{trend}		0.10		0.95	
<i>AGCGT</i> (Hap2)					
0	4/71	1.00	48/46	1.00	
1	8/88	2.39 (0.39–14.56)	55/81	0.62 (0.34–1.16)	0.48
2	2/20	2.87 (0.23–36.02)	32/21	1.80 (0.77–4.22)	
P_{trend}		0.33		0.47	
<i>TCTGG</i> (Hap3)					
0	8/147	1.00	117/120	1.00	
1	6/28	7.06 (1.21–41.15) ^c	16/28	0.54 (0.24–1.24)	0.04
2	0/4	NA	2/0	NA	
P_{trend}		0.05		0.28	
<i>TCTGG</i> (Hap3)					
0	8/147	1.00	117/120	1.00	
1 and 2 (grouped)	6/32	6.70 (1.17–40.63) ^c	18/28	0.59 (0.26–1.32)	0.02

^aNumber of cases/controls.

^bAdjusted for age at diagnosis (continuous), gender (male and female), ethnicity (Caucasian and non-Caucasian), consumption of fruit and vegetable (continuous), occupation risk (yes and no), and pre-existing lung disease (yes and no; exclusive of asthma); NA = not applicable because of small numbers.

^c $P_{\text{value}} < 0.05$.