Genotypes of glutathione transferase M1 and P1 and their significance for lung DNA adduct levels and cancer risk

David Ryberg1,2, Vidar Skaug1, Alan Hewer2, David H. Phillips2, Lorna W. Harries3, C. Roland Wolf3, Dagfinn Øgreid4, Arve Ulvik4, Phuong Vu1 and Aage Haugen1

1Department of Toxicology, National Institute of Occupational Health, P.O. Box 8149 Dep, N-0033 Oslo, Norway, 2Section of Molecular Carcinogenesis, Haddow Laboratories, Institute of Cancer Research, Cotswold Road, Sutton, Surrey SM2 5NG, UK, 3Molecular Pharmacology Unit, The Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee, DD1 5SY, UK and 4Department of Oncology, Haukeland Hospital, N-5021 Bergen, Norway

© Oxford University Press 1285

ACCELERATED PAPER

Introduction

Tobacco smoke contains several thousand chemicals of which about 50 compounds are known carcinogens. The most important carcinogens belong to the groups polycyclic aromatic hydrocarbons (PAHs*), aromatic amines and nitroso compounds. Rather than the parent compound itself, it is usually metabolites of the carcinogens that initiate cancer. Many of these compounds are oxidized by phase I enzymes, represented by cytochrome P450 enzymes, into reactive metabolites which are detoxified by phase II enzymes. One important class of the latter group is glutathione S-transferases (GSTs). These enzymes have an important role in protecting DNA against damage and adduct formation through glutathione conjugation to electrophilic substances, particularly those with lipophilic groups. Four different gene families of GSTs (cytosolic enzymes) are known, α, μ, π, and θ (1). Each family may comprise several genes; for the μ family at least five genes are described, but only one gene is known for the π family (1). The various enzymes have different but often overlapping substrate specificities (1). Also different patterns of expression in different cells and tissues are known (1–3).

Large interindividual variations in enzyme activities have been demonstrated for several GSTs (1,2). Some of these variations are genetically linked and may affect individual cancer risk. The GSTM1 gene is one of the most extensively studied concerning metabolic polymorphisms and cancer risk. About 50% of the Caucasian population is homozygous for a deletion in the gene which makes the enzyme inactive (null genotype) (4,5). The gene product is suggested to be particularly important for detoxifying epoxides of certain carcinogenic PAH compounds (1,6). Several case-control studies have indicated that the GSTM1 null genotype is associated with increased lung cancer risk (7). However, conflicting results are also published, which at least may be interpreted as indicating that the effect of this single genetic polymorphism on cancer risk is low. Other studies have indicated that presence of intact GSTM1 gene may be protective for cytogenetic damage (8) and carcinogen-derived DNA adduct formation (6,9,10). Our recent study on DNA-adduct levels in normal lung tissue from lung cancer support this conclusion (11). Since GSTM1 is only weakly expressed in lung tissue (1,3), the mechanisms of its protective role in this tissue is still insufficiently understood.

The major GST protein in human lung is GSTP1-1 (1,3). The enzyme is active towards many epoxides of PAHs (1,6). Thus it may in theory compensate for loss of GSTM1 activity. Increased expression of the GSTP1 gene has been found in tumors in several tissues including lung (1). The gene is also suggested to be involved in the development of acquired resistance towards anti-cancer drugs (1,12). A polymorphic site at codon 104 (A to G substitutions replacing isoleucine with valine) in the P1 gene is known which change the kinetic properties of the enzyme (13). In a recent case-control study the GG genotype was significantly more frequent among patients with bladder and testicular cancer, and the AA genotype was decreased in prostate cancer (14).

In this study of male cancer patients with NSCLC we have examined genotypes of GST M1 and P1 genotypes and their relation both to lung cancer risk and DNA adduct levels found in normal lung tissue.

Materials and methods

Samples from normal lung parenchyma and tumor tissue were obtained from 140 previously untreated male lung cancer patients undergoing surgery. A questionnaire was filled out for each patient where information on smoking habits were recorded. The mean age, smoking duration and smoking dose (with standard deviation) were 62.3 ± 10.3 years, 44 ± 10 years and 35 ± 17 packyears, respectively. All patients were current or former smokers. Histological classification of tumors was made according to the WHO guidelines.

© Oxford University Press 1285

*Abbreviations: GST, glutathione S-transferase; PAH, polycyclic aromatic hydrocarbons; NSCLC, non-small cell lung cancer; SSCP, single strand conformation polymorphism.
The two-dimensional thin layer chromatography of 32P-postlabelled DNA with one or two G alleles compared to those with high levels, had a genotype distribution very similar to the healthy controls. The comparison of means were performed by t-test and ANOVA.

Results

We have examined the A-G polymorphism at codon 104 in the GSTP1 gene in 138 male cancer patients with primary NSCLC. When the genotype distribution was compared with data from 297 healthy male controls (Table I), a statistically significant difference was found ($\chi^2 = 7.49, P = 0.024, 2 df$). This difference was due mainly to the homozygous genotypes AA and GG which contributed to ~90% of the total $\chi^2$ value. As indicated in Table I lung cancer patients had significantly lower incidence of the AA genotype and higher incidence of GG compared to controls. When the controls were matched with respect to age and lifetime smoking dose (see Materials and methods), similar results were obtained (113 control subjects). This indicated that individuals with the GSTP1 GG genotype may be more susceptible than those with AA. An odds ratio of 1.90 was obtained when the GG genotype was compared with AG and AA in cases and controls (Table I). When the patients were stratified according to the two major histological tumor types, it appeared that the described difference was found in the group of squamous cell carcinomas, and not among patients with adenocarcinomas (Table I). The lung cancer group was also stratified by age and smoking habits. Since the number of patients homozygous for the G allele were few, it became necessary to group the hetero- and homozygous. The highest frequency of patients with G alleles (AG and GG) was found among those below 50 years of age (13/15, 87%) which was significantly higher than the controls (Fisher exact test $P = 0.006$; odds ratio, 6.73; 95% confidence interval, 1.49–30.35). Similar results were obtained for patients smoking <30 years before the diagnosis (data not shown).

Concerning lifetime smoking dose (expressed as packyears) we found no particular patient groups with exceptional genotype distributions when patients with G alleles were grouped. The hydrophobic DNA adduct levels were determined in normal lung tissue from 70 current smoking lung cancer patients as described in Materials and methods. When the patients were grouped according to their GSTP1 genotype, significantly higher mean adduct levels were found in patients with one or two G alleles compared to those with the AA genotype (Table II). The highest DNA adduct level was found in the GG group, however, not statistically different from the AG group ($P = 0.39$). Analysis of variance in log transformed data indicated a linear trend in the adduct level with increasing number of G alleles (Table II). The patients were also ranked according to the lung adduct levels. The relationship between adduct level and GSTP1 genotype was also confirmed when the genotype distribution was analyzed in the upper and lower tertile adduct group (Table II). Patients with the lowest level of hydrophobic DNA adducts (lower tertile group), contrary to those with high levels, had a genotype distribution very similar to the healthy controls. The AA genotype of GSTP1 could therefore partially protect the DNA from being damaged by reactive hydrophobic (PAH) compounds.

We found an increased incidence of the GSTM1 null genotype among lung cancer patients; however, this was not statistically different from that found in 342 healthy controls (Figure 1). Matching the control group with respect to age and smoking dose (see Materials and methods) did not change this conclusion. The patients were also stratified by histology, age and smoking dose. The null genotype was most prevalent in squamous cell carcinomas and in patients smoking <20 packyears (lowest quartile group) and among those below 56 years of age (lowest quartile group) (Figure 1). When data on hydrophobic DNA adduct levels were analyzed in relation to the GSTM1 genotypes, a higher mean level was found in patients with the null genotype than among those with at least one intact GSTM1 allele (12.5 ± 8.5 and 9.4 ± 6.0 adducts per 10^8 nucleotides, respectively, $P = 0.088, T = 1.72$). This

---

**Table I. GSTP1 genotypes in male lung cancer patients and healthy controls**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Controls</th>
<th>Lung cancer patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. (%)</td>
<td>All patients No. (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Squamous cell No. (%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adeno carcinomas No. (%)</td>
</tr>
<tr>
<td>AA</td>
<td>153 (51.5)</td>
<td>53 (38.4)(^a)</td>
</tr>
<tr>
<td>AG</td>
<td>117 (39.4)</td>
<td>63 (45.7)</td>
</tr>
<tr>
<td>GG</td>
<td>27 (9.1)</td>
<td>22 (15.9)(^b)</td>
</tr>
</tbody>
</table>

\(^a\)AG + GG against AA in patients versus controls: 2 sided $P = 0.011$; odds ratio, 1.70; 95% confidence interval, 1.13 – 2.57.

\(^b\)AG against AG + AA in patients versus controls: 2 sided $P = 0.035$; odds ratio, 1.90; 95% confidence interval, 1.04 – 3.47.

\(^c\)AG + GG against AA in patients versus controls: 2 sided $P = 0.001$; odds ratio, 2.50; 95% confidence interval, 1.41 – 4.42.

\(^d\)AG against AG + AA in patients versus controls: 2 sided $P = 0.015$; odds ratio, 2.41; 95% confidence interval, 1.17 – 4.96.

(15). All except three tumors were NSCLC. The control group comprised 342 healthy males who are present or previous workers in several different companies in Norway. All individuals were anonymized after collection of data and blood samples. Only individuals born in Norway and of Norwegian citizenship were used. (The same ethnic background as the patients.) Mean age of the controls was 54 ± 15 years. In this group 35% were never-smokers. The mean smoking duration and dose for smokers were 32 ± 12 years and 21 ± 13 packyears. A group of controls was made which matched the average age and smoking dose of the patients. Selecting all subjects with age above 50 years who had smoked >22 packyears satisfied this requirement. This ‘matched’ group comprised 113 individuals with age 63.5 ± 8.3 years and smoking dose 35.3 ± 11.5 packyears. DNA was isolated from frozen tissue and blood (controls) using standard procedures with proteinase K/RNase digestion and phenol/chloroform extraction.

The GSTM1 genotype was determined in genomic DNA, using a PCR method as described previously (16). The A-G polymorphism at codon 104 in the GSTP1 gene was determined by PCR and digestion of the product with Alu261 as described by Harries et al. (14). One hundred and thirty-three patients and 297 controls were analyzed for both genotypes. DNA adducts were measured in normal lung tissues from 70 patients by 32P-postlabelling analysis with the nuclelease P1 modification (17). All these patients were current smokers. Data from 38 of these cases have been published previously (16). The two-dimensional thin layer chromatography of 32P-postlabelled DNA digests displayed diagonal zones characteristic of DNA adducts formed from complex mixtures of aromatic and/or hydrophobic compounds. Tumor DNA from 112 male patients was screened for mutations in exon 1 of the Ki-ras gene using PCR and a minigel SSCP method (18). Samples with aberrantly migrating bands were sequenced using Sequenase 2.0 and conditions described by the supplier (USB, Cleveland, OH). Previously published mutational data from exons 4–9 in the p53 gene were used from 83 male patients (19).

Chi-squared test and Fisher’s exact test were used in frequency analysis. Two-sided P-values are given when not indicated otherwise.
Table II. Hydrophobic DNA adduct levels normal lung tissue from male lung cancer patients in relation to GSTP1 genotypes

<table>
<thead>
<tr>
<th>GSTP1 genotype</th>
<th>Mean adduct level ± SD (N)</th>
<th>Number of patients with each genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adducts per 10⁸ nucleotides</td>
<td>Highest tertile adduct group</td>
</tr>
<tr>
<td>AA</td>
<td>7.89 ± 5.11&lt;sup&gt;a,b&lt;/sup&gt; (N = 25)</td>
<td>4 (17.4%)</td>
</tr>
<tr>
<td>AG</td>
<td>12.09 ± 7.44&lt;sup&gt;b&lt;/sup&gt; (N = 35)</td>
<td>14&lt;sup&gt;a&lt;/sup&gt; (60.9%)</td>
</tr>
<tr>
<td>GG</td>
<td>15.54 ± 10.23&lt;sup&gt;b&lt;/sup&gt; (N = 10)</td>
<td>5&lt;sup&gt;a&lt;/sup&gt; (21.7%)</td>
</tr>
</tbody>
</table>

Log transformed adduct data were analysed by ANOVA.

<sup>a</sup>AA versus AG: $t = -2.66$, df = 50; $P = 0.010$. AA versus GG: $t = -2.55$, df = 15; $P = 0.022$.  
<sup>b</sup>Linear trend test: $F = 6.92$, $P = 0.011$.  
<sup>c</sup>Fisher exact test of (AG + GG) against AA in patients versus controls (Table I): 2 sided $P = 0.002$; odds ratio, 5.05; 95% confidence interval, 1.68–12.20.

Fig. 1. Frequency of GSTM1 null genotype in male lung cancer patients analyzed in indicated subgroups. Hydrophobic DNA adducts were measured in normal lung tissue from 70 current smokers. The adduct levels were ranked according to size and the GSTM1 null genotype frequency determined in the upper 33 percentile (with highest adduct level) and in the lower 33 percentile. Mutations in the p53 (exons 4–9) and Ki-ras (exon 1–2) genes were analyzed in tumors from 83 and 112 male lung cancer patients, respectively. Patients were grouped according to the mutational types (transversions at G:C base pairs and transitions) found in either one or both genes. In cases where different types were found in the two genes, the patient was included in both groups (3 cases). The genotype distribution in groups represented by solid bars were compared to all controls. The results of the statistical analysis are given above the bars.

Discussion

Presently, the GSTM1 gene is probably the most well studied gene in relation to metabolic polymorphism and lung cancer risk, although conflicting results exist. In a recent meta-analysis of 12 case-control studies comprising total 1593 cases and 2135 controls it was concluded that GSTM1 deficiency was a moderate risk factor for lung cancer development with an odds
ratio of 1.41 (7). A similar risk was estimated in this study (odds ratio 1.33) although statistical significance was not present, probably because larger number of subjects are necessary for detecting such a low increase in cancer risk. It is suggested that the influence of genetic factors may be more pronounced among patients who contract lung cancer at younger age and after relatively lower smoking doses. In this study the highest frequency of null genotype was found in these groups. In individuals which lack the protective role of GSTM1 (null genotype) a greater fraction of the relevant carcinogens in the cigarette smoke may theoretically reach cellular DNA and form carcinogenic adducts. The examination of genotypes in relation to the level of hydrophobic DNA adducts in normal lung tissue from patients may support the moderate cancer susceptibility associated with the null genotype. There is no simple relationship between adduct levels and cancer risk. However, animal studies have indicated that high levels of stable DNA adducts in the target tissue are associated with increased probability that tumors will occur (22). The examination of GSTM1 genotypes in several stratified patients group usually gave only weak statistical significance for each group, however, the overall pattern in the data consistently indicated that the null genotype may be associated with a slightly increased risk. In the above analysis patients were compared to controls matched with respect to ethnic background and gender but not by age and smoking dose. The average age and smoking dose of the controls was lower than for the patients. About one third of the controls were also non-smokers. Since lung cancer is strongly related to long-lasting smoking, some of the controls could therefore be in the patient group if they had been older and smoked more and longer. Since the present analysis indicated higher frequency of null genotype among patients than in unmatched controls, this probably means that matching the controls would have given even larger differences. The genotype distribution in a control group which matched with the average age and smoking dose of patients was very similar with the whole control group. This is not in conflict with the low effect of the null genotype concerning lung cancer risk.

When we examined the codon 104 polymorphism of the GSTP1 gene in the same material, it appeared that patients had significantly lower frequency of the AA and higher frequency of the GG genotype than controls. Thus individuals homozygous for the G allele may be of increased cancer risk. This is also supported by the data from measurement of DNA adduct levels in lung tissue. Due to the low frequency of the GG genotype in the population, an extended study is necessary to get a reliable estimate of lung cancer risk for homozygous and heterozygous. A similar problem is present for the risk associated with the GSTM1 genotypes where distinction between individuals with one and two intact alleles may be difficult due to methodological reasons.

Both GSTM1 and P1 are known to catalyze detoxification of PAHs in vitro, however, they may be differently expressed in different tissues (1,3). The expression of GSTM1 is high in the liver, but the level in lung is very low. The opposite pattern is found for the P1 gene. Increased lung cancer risk associated with lack of GSTM1-1 activity may therefore be explained by an important role for the liver in the metabolism of compounds entered through the lung (6). According to these suggestions the lung tissue are exposed to carcinogens both from the lumen side and from the bloodstream by metabolites recirculated from the liver. The higher incidence of the susceptible genotypes of both GST M1 and P1 in patients with squamous cell carcinomas may be in accordance with other studies indicating a more crucial role for the GST metabolism in the more proximal parts of the lung where squamous cell carcinomas are usually located (3,6,23). Since several GSTM genes are localized in a cluster on chromosome 1p (24), it is also possible that the actual gene studied is not only M1 but also the effect of other genes of this family expressed in the lung tissue (23,25). Immunohistochemical data have revealed large interindividual variations in the expression of M1 gene in lung (3). There is also an indication that the activity or lack of activity of the GSTM1-1 may affect the expression (inducibility) of CYPIA1, probably through regulatory mechanisms (26). This CYP gene is essential for the metabolism of many carcinogenic PAH compounds. Lung cancer patients who were recent smokers had significantly induced aryl hydrocarbon hydroxylase (AHH) activities in lung parenchyma compared with smoking non-cancer patients (27,28). Concerning DNA adduct formation, the level of adducts in recent smokers seemed to be positively correlated to AHH (CYP1A1) activity in lung tissue (27,28).

Irrespective of mechanisms and genes involved, our data may indicate that the null genotype of GSTM1 has less penetrance as a risk factor for lung cancer than genotypes of GSTP1. However, also for the GSTP1 it remains to clarify the role of the A to G substitution at codon 104. This results in an isoleucine to valine substitution in the hydrophobic binding region in the protein, which seems to give only a moderate change in its catalytic properties (13). This polymorphism may therefore be linked to other features in the P1 gene. In contrast to GSTA1/A2 and GSTM1, the expression level of P1 in the lung seems to be rather constant between individuals (23,29).

There is some theoretical support that biotransformation of certain mutagens may be affected both by GSTM1 and GSTP1 enzymes. The significance of combined genotypes at these two genloci concerning DNA damage and cancer risk remains

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Controls</th>
<th>Lung cancer patients&lt;sup&gt;a,b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GSTM1 null No. (%)</td>
<td>GSTM1 positive No. (%)</td>
</tr>
<tr>
<td>GSTP1 AA</td>
<td>72 (24.2)</td>
<td>81 (27.3)</td>
</tr>
<tr>
<td>GSTP1 AG or GG</td>
<td>68 (22.9)</td>
<td>76 (25.6)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Homogeneity test of odds ratio in the case and control group: \( P = 0.39, \) common or = 1.12, 95% CI \((0.77 - 1.65)\).

<sup>b</sup>Distribution of combined genotypes in cases versus controls: \( \chi^2 = 10.84, df = 3, P = 0.013 \).

<sup>c</sup>GSTP1 (AG or GG) + GSTM1 null against all other genotype combinations in patients versus controls: 2 sided \( P = 0.003; \) odds ratio, 1.96; 95% confidence interval, 1.26 - 3.06.
to be clarified. Our data indicate that this is a relevant and important question.

Acknowledgements

We thank L. Stangeland and A. Naalsund for providing human tissue and patient data, and A. Myking for assisting in the histological typing. This work was supported by grants from the Norwegian Cancer Society, the UK Cancer Research Campaign and EC BIOMED contract no CT94-1608.

References


