Gene–environment interaction in tobacco-related cancers

Emanuela Taïoli*

Department of Epidemiology, University of Pittsburgh Cancer Institute,
UPMC Cancer Pavilion, 5150 Center Avenue, Pittsburgh, PA 15232, USA

"To whom correspondence should be addressed. Tel: +412 623 2217;
Fax: +412 623 3878; Email: taiolies@upmc.edu

This review summarizes the carcinogenic effects of tobacco smoke and the basis for interaction between tobacco smoke and genetic factors. Examples of published papers on gene–tobacco interaction and cancer risk are presented. The assessment of gene–environment interaction in tobacco-related cancers has been more complex than originally expected for several reasons, including the multiplicity of genes involved in tobacco metabolism, the numerous substrates metabolized by the relevant genes and the interaction of smoking with other metabolic pathways. Future studies on gene–environment interaction and cancer risk should include biomarkers of smoking dose, along with markers of quantitative historical exposure to tobacco. Epigenetic studies should be added to classic genetic analyses, in order to better understand gene–environmental interaction and individual susceptibility. Other metabolic pathways in competition with tobacco genetic metabolism/repair should be incorporated in epidemiological studies to generate a more complete picture of individual cancer risk associated with environmental exposure to carcinogens.

Tobacco carcinogens

Tobacco smoke is the most widespread carcinogen in the world. More than 3000 chemicals have been isolated from processed tobacco leaves (1). These are not only leaf constituents but also products derived from the soil, the atmosphere, the use of agricultural chemicals and from the process of curing, casings and flavoring of the leaves.

In addition to toxic compounds, unburned tobacco products contain carcinogenic nitrosamines derived from nitrates, amines, proteins and alkaloids present in the leaves; polycyclic aromatic hydrocarbons resulting from the processing; radioactive elements absorbed from the soil and the air and cadmium from cadmium-rich soils.

When tobacco is burned during smoking, many other reaction products are formed, among which are > 4000 identified chemicals and an unknown number of unidentified chemicals (1).

The products of mainstream smoke can be divided into particulate and gas phases. The particulate phase contains nicotine, nitrosamines [4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone], N-nitrosornornicotine, metals (cadium, nickel, zinc and polonium-210), polycyclic hydrocarbons and carcinogenic amines (4-aminobiphenyl). The vapor phase contains among the others carbon monoxide, carbon dioxide, benzene, ammonia, formaldehyde, hydrogen cyanide, N-nitrosodimethylamine and N-nitrosodibutylamine (2).

Approximately 60 known carcinogens are present in tobacco smoke (3), the strongest are PAHs, N-nitrosamines and aromatic amines and the most prevalent present in the vapor phase are aldehydes, benzene and butadiene (4).

Tobacco smoke and cancer

Tobacco from both active and passive smoking has the main point of entry into the body via the airways; some constituents dissolve in saliva and are absorbed or swallowed. Alcoholic drinks act as solvents of the smoke constituents, thus facilitating their absorption. Virtually, all the organs and tissues are reached by the active products of smoking. Data from epidemiological studies confirm the widespread action of tobacco smoke on tissues and organs (3).

In fact, in addition to the well-known link between smoking and lung cancer, large epidemiological studies have shown an association of smoking with several other cancer sites, either because of direct contact of the smoking products with the specific organ or because the organ is reached by smoking carcinogenic products through the blood stream. The long list of tobacco-associated cancers includes nose, oral cavity, oropharynx, hypopharynx, larynx, esophagus, pancreas, bladder, kidney, stomach, liver, colon, cervix and myeloid leukemia (5).

Much is known by now about the mechanisms by which carcinogens present in tobacco smoke act as both initiators and promoters of cancer in these organs and tissues. Several biomarkers of exposure, internal dose and effect have been developed and validated (Figure 1).

Genetic pathways of tobacco metabolism

PAHs, N-nitrosamines and aromatic amines are metabolized by a two-phase process. Phase I involves the activation of the carcinogen by enzymes encoded by the CYP gene superfamily. These enzymes are involved in the oxidative metabolism of several exogenous compounds, drugs and endogenous hormones. Cytochrome p450 1A1 is responsible for the first step of PAH metabolism. Other enzymes, such as CYP2C9, CYP1B1 and CYP2D6, are responsible for the activation of benzo[a]-pyrene and nitrosamine 4-(methylnitrosamine)-1-(3-pyridyl)-1-butanone, whereas CYP2E1 metabolizes 4-(methylnitrosamine)-1-(3-pyridyl)-1-butanone.

During the Phase II process, carcinogens are transformed into hydrophilic elements to facilitate excretion. Glutathione S-transferases are mostly responsible for this process. This multigene superfamily includes four classes α, μ, θ and π and detoxifies carcinogens from cigarette smoke as well as from other sources.

DNA damage is usually repaired by a series of genes specialized in this activity. Several polymorphisms in DNA repair genes have been identified, but their impact on repair phenotype and cancer susceptibility remains uncertain (6).

Since Phase I enzymes induce the formation of active carcinogens from procarcinogens, whereas Phase II enzymes conjugate these compounds and make them suitable for excretion, it is reasonable to think that the overall carcinogenic effect of tobacco compounds should be measured as the final result of the combined action of the two categories of enzymes. One central issue when looking at metabolic pathways is that the number of subjects who carry the ‘unfavorable’ gene polymorphism combinations may be very small, thus challenging the epidemiologists who have to assemble a large study in order to have enough subjects available with that specific genetic profile.

The expected frequency of combined metabolic gene polymorphisms can be calculated from the known frequencies of each polymorphism in the general population. Using this method, for example, the expected frequency of CYP1A1 msp1 variant + glutathione S-transferase μ (GSTM) deletion + glutathione S-transferase θ (GSTM1) deletion + glutathione S-transferase θ (GSTM1) variant + epoxide-hydrolase3 and 4, all genes involved in the metabolism of PAHs, turns out to be ~7 × 10⁻³ or 0.7 subjects in 10 000 carrying this particular genetic combination. Since we anticipate that not all the subjects will be exposed to the environmental factor of interest, in this case tobacco smoking, we can predict that the size of any study looking at this specific genetic pathway–environment interaction must be very large.

Genetic susceptibility and markers of DNA damage

Most of the metabolic and DNA repair genes carry polymorphisms that are present in the general population at various frequencies. Some
of these genetic variations alter the original gene function, thus increasing or decreasing the activity of the corresponding enzyme. For example, both GSTM1 and GSTT1 genes can be deleted and thus their conjugating activity can be absent. GSTM1 homozygous deletion is present at frequencies that vary from 30 to 50% in the general population, whereas the deletion of GSTT1 is around 20–30% in healthy subjects (7). Changes in enzymatic activity associated with polymorphisms in these genes may play a significant role in tobacco-related cancer risk and genetic susceptibility.

Most of the genetic polymorphisms described in the literature vary in frequency across ethnicity and geographic areas (7). This may be a further contributing factor to the observed variation in tobacco-related cancer incidence among smokers with different ethnic background.

Epidemiologists have employed several strategies to study individual genetic susceptibility to tobacco carcinogens. One possibility is to assess whether subjects carrying a set of polymorphisms in metabolic and DNA repair genes are at higher risk of tobacco-related cancers when compared with subjects exposed to the same amount of tobacco smoke, but carrying a more favorable metabolic/repair genetic profile. The expectation is that the same level of environmental exposure to a carcinogen will result in more damage in subjects who do not metabolize and/or repair well the genotoxic effect induced by the carcinogen. This approach has been used very extensively, with conflicting results.

Another appropriate strategy is to look at an intermediate end point of genotoxicity. A very commonly used measure of DNA damage induced by tobacco is DNA adducts. It is expected that adduct levels differ significantly with both case–control and exposure status, being higher in cancer cases, and among smokers.

Several epidemiologic studies have reported associations between DNA adduct levels and the most prevalent tobacco-related cancers including cancer of the lung, head and neck and bladder. A comprehensive meta-analysis confirmed this association in current smokers only (8), whereas for other cancers this evidence is less convincing. It is unclear how the measurement of DNA damage (adducts) in subjects already diagnosed with cancer can be used as an etiologic proof of carcinogenicity rather than an effect of the cancer itself.

Variations in genes involved in tobacco metabolism and/or DNA repair should produce a difference in local carcinogen levels; therefore, changes in levels of DNA damage should be observed as a consequence of the polymorphisms. Several studies have been conducted on changes in DNA adduct levels with polymorphisms in CYPs, GSTs and NATs genes in peripheral lymphocytes, lung and other target tissues, with conflicting results (9).

DNA damage in healthy target tissues

A further step forward is to consider DNA damage in target tissues, with the aim of assessing the effects of carcinogen metabolism/DNA repair where the carcinogen process starts, rather than in white blood cells, a peripheral tissue that reflects and funnels a multiplicity of effects, from environmental exposure to tissues’ local metabolisms, to individual susceptibility.

A summary of the literature that addresses DNA adducts in several target tissues, tobacco smoking and the influence of several metabolic gene polymorphisms is reported in Table I. Selection has been made so that only studies reporting DNA adducts in target tissues from healthy subjects or in healthy tissue from cancer patients are included in the table. A total of 28 distinct studies were identified in the literature; 14 of them were conducted on lung tissue from subjects with non-malignant diseases or from healthy areas of lungs from subjects with cancer. The most common gene polymorphism studied in conjunction with DNA adduct levels and smoking exposure was GSTM1 (22 studies), followed by CYP1A1 (12 studies) and GSTP1 (10 studies).

Despite the wealth of information produced in the last decade, no clear pattern emerges from the published studies; GSTM1 deletion appears to be associated with DNA adducts in lung tissue in 8 of the 14 studies conducted on this target tissue, whereas 5 of them did not find any association and 1 did not test this particular gene.

More recent studies included the analysis of ‘high-risk’ genotype combinations and showed that the presence of more than one polymorphism on the pathway to tobacco metabolism may act as effect modifiers of the role of tobacco in determining DNA damage in relevant target tissues. For example, 5 of the 14 studies conducted on lung tissue examined the combined effect of several polymorphisms on DNA adducts and reported significant association in the expected direction in all instances.

Data on other healthy target tissues are sparse and the results less clear-cut: five studies included healthy breast tissue/epithelial ductal cells. In three instances, NAT2 was tested in relation to smoking, and an association with adducts was observed in two studies. A significant effect of the combination of the myeloperoxidase and GSTP1 polymorphisms on DNA adducts was reported in one study.

The discrepancies between the results observed when studying markers of genotoxicity and metabolic gene polymorphisms in target tissues do not follow the pattern expected from what is known about the gene function and the tissue metabolic activity. This may be attributed to several factors, for example, a less than optimal definition of exposure (smoking status, intensity and duration), the lack of specificity and/or the high variability of the measure of DNA damage (DNA adducts), the difference in detection limits among laboratory techniques used for DNA adduct measure, the complex genetic pathways involved, which are only partially explored when looking at few genetic polymorphisms, and the presence of multiple different competing substrates for the same genetic pathway. In addition, it is possible that the metabolism of a target healthy tissue from a cancer patient is affected by the neighborhood presence of a cancer, thus making the correct interpretation of some of the results challenging.

An important aspect that is rarely addressed is inter- and intra-subject variability of the measures, and the comparison between markers measured in a target tissue and in peripheral blood. Time of persistence of the relevant biomarker in the target tissue and in peripheral blood also needs to be addressed properly.

![Table I. Examples of biomarkers currently used to study the pathway from exposure to tobacco smoke to cancer.](image-url)
<table>
<thead>
<tr>
<th>Author</th>
<th>Population</th>
<th>Laboratory tests</th>
<th>Tissue</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shields (10)</td>
<td>Mixed Caucasian and AA</td>
<td>32P-postlabeling PAH–DNA adducts—immunoaffinity chromatography, serum cotinine, CYP1A1 exon 7, GSTM1</td>
<td>38 lung tissues, 7 positive for adducts</td>
<td>No relationship between adducts and cotinine, gender or CYP1A1. Significant association adducts/GSTM1 deletion, confirmed by multivariate analysis.</td>
</tr>
<tr>
<td>Ryberg et al. (11)</td>
<td>Caucasian (Finn and Norwegian)</td>
<td>32P-postlabeling DNA adducts, GSTM1 and GSTP1</td>
<td>Healthy lung tissue from 63 previously untreated lung cancers</td>
<td>Adducts significantly higher in smokers than in non-smokers; weak association between smoking dose and adducts. Smoking adjusted, females had higher adduct levels than males. Higher adducts with GSTM1 deletion.</td>
</tr>
<tr>
<td>Budawi et al. (12)</td>
<td>USA (mixed)</td>
<td>32P-postlabeling DNA adducts, NAT1 and NAT2</td>
<td>Human urinary bladder mucosa (n = 26)</td>
<td>Association between adducts and NAT1. No data on smoking. Highest adducts with NAT1/NAT2 combination.</td>
</tr>
<tr>
<td>Kato et al. (13)</td>
<td>USA (mixed)</td>
<td>32P-postlabeling 7-methyl-d-GMP and PAH-d-GMP DNA adducts, CYP2D6, CYP2E1, GSTM1, CYP1A1, cotinine</td>
<td>95 healthy lungs from autopsies</td>
<td>Association between 7-methyl-d-GMP adducts and CYP2D6, CYP2E1, restricted to non-smokers. PAH-d-GMP DNA adducts associated with GSTM1 deletion.</td>
</tr>
<tr>
<td>Ryberg et al. (14)</td>
<td>Caucasian (Finn and Norwegian)</td>
<td>32P-postlabeling DNA adducts with nuclease P1 modification, GSTM1 and GSTP1</td>
<td>70 non-tumor lung tissue from lung cancers, current smokers</td>
<td>Adducts significantly associated with GSTP1 and GSTM1. GSTM1/ GSTP1 AA = 6.68 ± 3.57; GSTM1/ GSTP1 GG = 19.03 ± 11.04 adducts x 10⁶ nucleotides.</td>
</tr>
<tr>
<td>Kadlubar et al. (15); Thompson et al. (16)</td>
<td>USA</td>
<td>OH-dG by HPLC, other oxidative stress adducts, GSTM1, T1, NQO1; 4-aminobiphenyl malondialdehyde adducts, NAT1, GSTM1, T1, P1</td>
<td>30 pancreatic tissue from organ donors</td>
<td>Adducts not associated with smoking or genotype; ABP associated with NAT1 genotype; no effect of smoking.</td>
</tr>
<tr>
<td>Schoket et al. (17)</td>
<td>Caucasian (Hungary)</td>
<td>32P-postlabeling DNA adducts with nuclease P1 modification, GSTM1, CYP1A1 MspI</td>
<td>70 non-tumor lung tissues from lung cancer, 26 healthy lungs</td>
<td>Significant association between adducts and smoking, no association with genotype.</td>
</tr>
<tr>
<td>Rojas et al. (18)</td>
<td>Caucasian (Russia)</td>
<td>BPDE–DNA adducts—HPLC, GSTM1, CYP1A1</td>
<td>20 non-tumor lung tissues from lung cancer</td>
<td>Significant association between adducts and GSTM1 deletion. Possible interaction between GSTM1 deletion and CYP1A1 2A/2A and 2A/2B on DNA adducts levels (only two subjects).</td>
</tr>
<tr>
<td>Whyatt et al. (19)</td>
<td>Caucasian (Poland)</td>
<td>Plasma cotinine, CYP1A1 MspI, BPDE–DNA adducts by ELISA</td>
<td>Placenta, 160 newborn with different exposures to PAH</td>
<td>No association between adducts and smoking in the mothers or cotinine levels. Adducts associated with CYP1A1.</td>
</tr>
<tr>
<td>Pfau et al. (20)</td>
<td>USA</td>
<td>32P-postlabeling DNA adducts, NAT1, NAT2</td>
<td>42 reduction mammoplasty</td>
<td>Association between adducts and NAT2, no effect of smoking.</td>
</tr>
<tr>
<td>Author</td>
<td>Population</td>
<td>Laboratory tests</td>
<td>Tissue</td>
<td>Results</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------</td>
<td>------------------</td>
<td>---------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Van Schooten et al. (21)</td>
<td>Caucasian (Netherland)</td>
<td>32P-postlabeling aromatic DNA adducts, GSTM1, GSTT1 and VDR genes</td>
<td>41 right atrial appendage</td>
<td>DNA adducts significantly elevated in smokers compared with ex-smokers or non-smokers; linear relationship with smoking dose. No association between adducts and genotypes.</td>
</tr>
<tr>
<td>Butkiewicz et al. (22)</td>
<td>Caucasian (Poland)</td>
<td>32P-postlabeling DNA adducts, CYP1A1, CYP2D6, GSTM1, GSTP1</td>
<td>165 non-tumor tissues from lung cancer</td>
<td>Significant association between smoking and adducts. Significant association between adducts and combination GSTM1 deletion/CYP1A1 Ile/Val.</td>
</tr>
<tr>
<td>Cheng et al. (23)</td>
<td>Asian (Taiwan)</td>
<td>32P-postlabeling aromatic DNA adducts, CYP1A1 MspI, GSTM1</td>
<td>Non-tumor lung tissues from 73 lung cancers, 33 non-cancers</td>
<td>In non-cancer tissues, no association between adducts and smoking or genetic polymorphisms.</td>
</tr>
<tr>
<td>Hardie et al. (24)</td>
<td>Caucasian</td>
<td>HPLC 8-OH-dG DNA adducts, hOGG1 polymorphism</td>
<td>37 paired lung non-tumor and tumor specimen</td>
<td>No association between adducts and genotype. No analysis of smoking data.</td>
</tr>
<tr>
<td>Schoket et al. (25)</td>
<td>Caucasian (Hungary)</td>
<td>32P-postlabeling DNA adducts, BPDE–DNA adducts by ELISA, GSTM1, CYP1A1, GSTP1 ile105Val, CYP1B1 Leu432Val, NQO1 Pro187Ser</td>
<td>94 non-tumor lung tissue from lung cancers, current and ex-smokers</td>
<td>No association between adducts and individual polymorphisms or with smoking dose. Interaction between GSTM1 and GSTP1 on adducts levels.</td>
</tr>
<tr>
<td>Li et al. (26)</td>
<td>USA</td>
<td>32P-postlabeling with nuclease P1 modification, 8-OH-dG by HPLC, CYP1A1, CYP2E1, NAT1, NAT2, GSTM1, MnSOD, hOGG1</td>
<td>Non-tumor pancreatic tissue from 13 cancers, 24 organ donors, 6 non-pancreatic cancers, 5 chronic pancreatitis</td>
<td>Association of DNA adducts with CYP1A1 polymorphism.</td>
</tr>
<tr>
<td>Chen et al. (27)</td>
<td>Asian (China)</td>
<td>GSTM1 and GSTP1, PAH–DNA adduct immunostaining; AFB1–DNA adducts and 4-ABP–DNA adducts—immuno-peroxidase methods</td>
<td>27 liver tissues from non-HHC patients; 99 non-tumor tissues from HCC patients</td>
<td>No differences in adducts with smoking or GSTM1 or GSTP1.</td>
</tr>
<tr>
<td>Brockstedt et al. (28)</td>
<td>Caucasian (Germany)</td>
<td>32P-postlabeling, HPLC DNA adducts; CYP1A1, CYP1A2, MPO, NQO1, GSTP1, GSTM1, GSTT1</td>
<td>26 breast tissue from reduction mammoplasty, 9 breast cancer</td>
<td>Significantly higher adducts with the MPO A-463 variant and with GSTP1'B or GSTP1'C (ns). Significantly higher adducts with MPO and GSTP1 combined variants than with all other genotype combinations.</td>
</tr>
<tr>
<td>Thompson et al. (29)</td>
<td>USA</td>
<td>32P-postlabeling DNA adducts, NAT2</td>
<td>50 epithelial ductal cells</td>
<td>66% of specimens show adducts, no association with NAT2.</td>
</tr>
<tr>
<td>Lewis et al. (30)</td>
<td>Caucasian (UK)</td>
<td>HPLC and P32-postlabeling N7-methylguanine DNA adducts, NQO1, GSTM1, GSTT, GSTP1, CYP2D6, CYP2E1</td>
<td>Bronchial lavage from 31 non-malignant patients, 13 lung cancers</td>
<td>Adducts significantly higher in smokers than non-smokers healthy subjects. Adduct higher with GSTM1 null, GSTT1 null or GSTP1 Ile/Ile genotypes, and in individuals with three unfavorable genotypes (GSTM1 null/ GSTT1 null and GSTP1 Ile/Ile) compared with others.</td>
</tr>
</tbody>
</table>
Another limitation of the available scientific literature is the variety of genes tested in each study, which are rarely comparable across studies.

Gene–environment interaction models

The theoretical framework of gene–environmental interaction has been clearly set a while ago by several authors (41–43). In the case of tobacco-related cancers, the model assumes that there is an association between the cancer and the environmental agent (in this case tobacco smoking), and in the absence of exposure, the presence or absence of the genetic risk factors is irrelevant for disease causation. However, in the presence of exposure, genetic polymorphisms may modulate the association observed between exposure and cancer.

It has been suggested that the degree of interaction between metabolic genes and tobacco smoke is not linear, but varies with exposure dose, thus making the picture more complex (44). For some genes, a greater degree of gene–environment interaction appears at lower doses of exposure (the interaction follows an inverse dose function), whereas for other genes a converse high-exposure gene effect is observed with the magnitude of interaction that increases as a function of dose.

Such phenomenon has been observed using different intermediate end points, such as micronuclei, DNA adducts and sister chromatid exchange. Case–control studies where smoking dose was available

<table>
<thead>
<tr>
<th>Author</th>
<th>Population</th>
<th>Laboratory tests</th>
<th>Tissue</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tang et al. (31); Tang et al. (32); Rundle et al. (33)</td>
<td>USA (mixed)</td>
<td>PAH–DNA adducts, immunohistochemistry, SULT1A1, XPD, GSTM1</td>
<td>Breast tissue from cases (n = 119) and benign breast disease (n = 108)</td>
<td>No association between adducts and SULT1A1 or XPD in benign tissue. No information on smoking; no association between adducts in benign tissue and GSTM1; no information on smoking.</td>
</tr>
<tr>
<td>Piipari et al. (34)</td>
<td>Caucasian (Finland)</td>
<td>GSTM1 and GSTP1, 32P-postlabeling DNA adducts</td>
<td>Bronchoalveolar macrophages of 31 smokers</td>
<td>No correlation between GSTM1 or GSTP1 genotypes and the aromatic DNA adducts.</td>
</tr>
<tr>
<td>Peluso et al. (35)</td>
<td>Caucasian (Italy)</td>
<td>32P-postlabeling, CYP1A1, GSTM1, GSTT1</td>
<td>55 nasal brushing and bronchoscopy from 32 healthy and 19 cancer patients, 4 unknown</td>
<td>Adducts in nasal mucosa correlated with smoking. No association with polymorphisms.</td>
</tr>
<tr>
<td>Van Schooten et al. (36)</td>
<td>Netherlands (Caucasian)</td>
<td>Aromatic DNA adducts by 32P-postlabeling, MPO –463G-A</td>
<td>33 smoking non-malignant patients, bronchoalveolar lavage</td>
<td>Significant association between DNA adducts and MPO polymorphism.</td>
</tr>
<tr>
<td>Paracchini et al. (37)</td>
<td>Caucasian (Italy)</td>
<td>PAH–DNA immunofluorescence, GSTM1</td>
<td>182 sperms from healthy men from infertility clinic</td>
<td>Adducts significantly higher in subjects with GSTM1− than in subjects with GSTM1+. Interaction with occupational exposure, but not with smoking.</td>
</tr>
<tr>
<td>Munnia et al. (38)</td>
<td>Caucasian (Italy)</td>
<td>32P-postlabeling chromatography MDA–DNA adducts, GSTM1, GSTT1, cyclin D G870A, urine thyocianate</td>
<td>43 subjects undergoing a bronchoscopic (26 healthy)</td>
<td>MDA–DNA adducts significantly higher in current smokers than in never smokers; higher levels of MDA–DNA adducts in subjects with GA and AA cyclin D1 polymorphism.</td>
</tr>
<tr>
<td>Nock et al. (39)</td>
<td>USA Caucasian (2/3) and AA</td>
<td>Immunohistochemical assays for PAH–DNA adducts, mEH His 139Arg and Tyr113His, CYP1B1 Leu432Val and Ala19Ser, CYP1A1 Ile462Val, GSTP1 Ile105Val, GSTM1</td>
<td>400 non-tumor prostate tissue from cancer patients</td>
<td>After adjusting for smoking status, the ‘high-risk’ genotype combination (CYP1B1 432Val/Val, mEH 139Arg/Arg, GSTP1 105Ile/Ile) was associated with increased adducts only in Caucasian non-tumor cells (0.2363 versus 0.1920 absorbance units; P = 0.05)</td>
</tr>
<tr>
<td>Ambrosone et al. (40)</td>
<td>USA</td>
<td>32P-postlabeling ABP and PhIP–DNA adducts, GSTs CYP1A1, 1A2, SULT1A1, NAT2, NAT1</td>
<td>64 samples, epithelial ductal cells</td>
<td>PhIP and ABP adducts associated with the rapid NAT2 and/or rapid NAT1 genotypes; no association with smoking.</td>
</tr>
</tbody>
</table>

HPLC, high-performance liquid chromatography; BPDE, benzo[a]pyrene diol-epoxide; VDR, vitamin D receptor; SULT1A1, Sulfotransferase; HDA, malondialdehyde.
confirmed that the strength of the gene–disease association varies across smoking dose.

**Issues to be considered when studying gene–environment interaction**

Although some of the complex methodological aspects of gene–environment interaction have been touched in the previous section, there are more factors to be considered when interpreting the results obtained by an epidemiological study on tobacco smoke, individual genetic factors and cancer risk.

One critical factor is that more details on smoking history, such as age at start smoking, smoking amount, duration and number of quitting attempts, may be important determinants in the interaction between smoking and genetic background, thus the relative weights of each factor should be taken into account (45).

For example, age at starting smoking may influence breast cancer, mostly because the immature breast tissue is more sensitive to tobacco genotoxicity. An unfavorable metabolic/DNA repair genotype may modulate breast cancer risk differently in women who started smoking early in life in comparison with others who started smoking later. The interaction between smoking and estrogen levels may play a significant role on such risk, and the degree of interaction may vary according to the age at starting smoking in relation to age at puberty.

The direct carcinogenicity of smoking on target tissues has been well documented by several experimental studies; however, the indirect activity of smoking on metabolic pathways related to other cancer risk factors is less easy to pinpoint in a human study.

For example, it is known that tobacco smoking inhibits the aromatization of androgens into estrogens, thus interacting with another metabolic pathway that is relevant to hormone-related and hormone-sensitive cancers. However, when conducting an epidemiological study, such additional gene–environment interaction is not usually considered.

The function of both metabolic and DNA repair genes may change substantially even in the absence of polymorphisms, and this may happen because of epigenetic changes induced by environmental factors.

Together with tumor suppressor genes methylation, both metabolic (CYP1A1 and GSTP1) and DNA repair genes (HMLH1 and MGMT) have been shown to be hypermethylated in several cancers, including lung cancer (46). The epigenetic phenomena were associated with smoking exposure, although the effect of environmental exposures on methylation pattern is still poorly studied.

A significant association between methylation of several genes and smoking has been reported (47,48); in a sporadic case, methylation of the p16 gene promoter was associated with starting smoking at a younger age (47). This information underlines the connection between environmental exposure and epigenetic events. Prospective studies on the predictive role of methylation of critical genes are needed in order to understand the public health relevance of these genetic markers as tools for identifying the early signs of genetic damage deriving from environmental exposure to carcinogens such as smoking. Gene methylation studies involving dose–response are also needed.

**Examples from epidemiological studies**

A large body of work has been conducted on the association between single nucleotide polymorphisms and cancer in relation to smoking habits. Several summary meta-analyses have been conducted with the purpose of understating the interaction between genetic factors, general measures of tobacco exposure such as ever versus never smokers and cancer. However, the subtle interaction between smoking dose/quantity and genetic factors on cancer risk cannot be fully elucidated by the reanalysis of published and existing large data sets (49). A review of several meta- and pooled analyses on the association observed between several Phase I and Phase II metabolic gene polymorphisms and lung cancer can be found in Schwartz et al. (50). Polymorphisms in CYP1A1 and combinations of Phase I and Phase II polymorphisms were found to be significantly associated with lung cancer. Polymorphisms in DNA repair genes were studied with less frequency, but the findings consistently point to a significant association with lung cancer.

Another interesting example is bladder cancer, a disease associated to both smoking and occupational exposure. The most studied polymorphisms in bladder cancer are N-acetyltransferases and glutathione S-transferases. A meta-analysis of bladder cancer studies and polymorphisms in these categories of genes showed that GSTM1 deletion and NAT2 slow acetylator status were the only two factors significantly associated with bladder cancer. An interaction with smoking was only observed for the NAT2 polymorphisms (51).

Data on other cancer sites, such as head and neck and colon, in conjunction with genetic polymorphisms on the pathway to smoking metabolism/DNA repair are less frequent and inconsistent in their results.

**Data from the GSEC study**

Summary data from a large ongoing pooled analysis of individual data Genetic Susceptibility to Environmental Carcinogens (GSEC study) support the concept that a clear demonstration of gene–tobacco smoking interaction is hard to obtain even by mining large epidemiological databases.

A pooled analysis of 34 case–control studies on GSTT1 and lung cancer, including 7044 cases and 10 000 controls (52), did not show statistical evidence of multiplicative interaction between GSTT1 and smoking, at least for Caucasians.

Another pooled analysis included on nine published and two unpublished case–control studies from the GSEC database on head and neck cancer, including 2334 cases and 2766 controls (53), and reported no differences in the magnitude of the association between head and neck cancer and smoking status according to GSTM1, GSTT1, GSTP1 and CYP1A1 polymorphisms.

Smits et al. (54) reanalyzed six studies on colorectal cancer, including 1130 cases and 2519 controls, and restricted the analyses to Caucasians. There was no interaction between the effects of smoking and GSTM1 genotype on colorectal cancer risk.

Benhamou et al. (55) pooled data from 9500 subjects involved in 21 case–control studies on lung cancer. No evidence of interaction between GSTM1 genotype and either smoking status or cumulative tobacco consumption on lung cancer risk was observed.

No interaction between GSTP1 and CYP1A1 on lung cancer risk was also reported (56), whereas the interaction was observed when the CYP1A1 exon 7 polymorphism was studied (57).

A reanalysis of 10 studies (1496 cases and 1444 controls) on bladder cancer and GSTM1-null genotype (58) indicated no evidence of multiplicative interaction between the GSTM1-null genotype and ever smoking in relation to bladder cancer, although there was a suggestion of additive interaction.

Another analysis on bladder cancer considered the role of NAT2 genotype (59) and included 1530 cases and 731 controls (all Caucasians). The risk of cancer was elevated in smokers and occupationally exposed subjects, with the highest risk among slow acetylators. The increase in risk was limited to current smokers (odds ratio = 1.74, 95% confidence interval = 0.96–3.15).

A pooled analysis from the GSEC database included 10 studies (3688 cases and 3874 controls) on lung cancer and myeloperoxidase G-463A polymorphism (60); the results showed an inverse association between the gene polymorphism and lung cancer. The myeloperoxidase G-463A polymorphism was significantly protective in ‘ever’ smokers but not in ‘never’ smokers.

Recently, we have analyzed the cumulative effect of variants in three metabolic genes, CYP1A1, GSTM1 and GSTT1, on lung cancercancer. The odds ratio for lung cancer was 8.25 (95% confidence interval = 2.29–29.77) for the combination CYP1A1*4 and the double deletion of both GSTM1 and GSTT1. The association was more pronounced in never smokers than ever smokers. The group of subjects carrying the three genes combination, however, represented a very small fraction of the entire population, thus limiting the public health value of such finding (61).
Conclusion

The assessment of gene–environment interaction in tobacco-related cancers has been more complex than originally expected for several reasons, including the multiplicity of genes involved in tobacco metabolism, the numerous substrates metabolized by the relevant genes and the interaction of smoking with other metabolic pathways.

Lack of appropriate study design and/or statistical power is also a cofactor. Assessment of details on smoking history, quantity and type has also proven to be hard to perform.

Future directions

Smoking is the main risk factor for a variety of cancer sites, and preventive measures against smoking initiation are still considered the best tools in the hands of epidemiologists. Smokers, however, still represent a sizable fraction of the population, and this is of great concern for health professionals because it is known that smoking cessation attempts are less successful than desired, mostly because of the addictive attributes of tobacco.

It is still important to detect genetic factors that interact with tobacco exposure and modify the individual risk of tobacco-related cancers. This approach may help predicting who, among smokers, will be more likely to develop a tobacco-related cancer in the short term.

Biomarkers of smoking dose should be incorporated in epidemiological studies on gene–environment interaction and cancer risk, along with markers of quantitative historical exposure to tobacco. Some of these tools still need to be developed. Epigenetic studies should be added to classic genetic analyses, in order to better understand gene–environmental interaction and individual susceptibility.

With the wealth of data generated by microarray studies, there is a need for the development and refinement of statistical methods to evaluate gene–environment interaction and effectively translate the laboratory data into public health strategies that can make an impact on cancer incidence.

Other metabolic pathways in competition with tobacco genetic metabolism/repair should be incorporated in epidemiological studies to generate a more complete picture of individual cancer risk associated with environmental exposure to carcinogens.

From an epidemiological point of view, it is still necessary to assemble large data sets and large well-characterized populations in order to test complex gene–environment interaction pathways. Biomarkers of DNA damage and exposure, along with susceptibility markers, should be incorporated in large, prospective cohort studies, in order to understand the temporal relationship between exposure, damage and cancer development within a well-designed epidemiological framework.

Funding

NIH SPORE lung cancer (2P50CA090440); ECNIS Network of Excellence, (EU contract 513943).

Acknowledgements

Conflict of Interest Statement: None declared.

References


Received October 22, 2007; revised February 27, 2008; accepted February 27, 2008