

Lung cancer susceptibility

Steen Mollerup¹, David Ryberg, Vidar Skaug and Aage Haugen

Department of Toxicology, National Institute of Occupational Health, P.O. Box 8149 Dep., N-0033 Oslo, Norway

1) To whom correspondence should be addressed, E-mail: Steen.Mollerup@stami.no

ABSTRACT

Worldwide, lung cancer is the most common malignancy and the leading cause of cancer mortality. The molecular epidemiology of this disease has received widespread attention, since the role of tobacco smoking in the etiology of lung cancer is well established. However, only about 15% of smokers develop lung cancer, indicating that susceptibility factors are involved. This review summarizes results we have obtained in an effort to understand gene-environment interactions that determine individual lung cancer risk. The focus has been on genes involved in carcinogen metabolism and sex differences in lung cancer risk, and on prognostic implementations.

Key words: lung cancer, susceptibility, sex, polycyclic aromatic hydrocarbons, cytochrome P450, glutathione *S*-transferase, p53

Abbreviations: AhR, aryl hydrocarbon receptor; CI, confidence interval; CYP, cytochrome P450; ER, estrogen receptor; GST, glutathione *S*-transferase; HR, hazard ratio; NSCLC, non-small cell lung carcinoma; OR, odds ratio; PAH, polycyclic aromatic hydrocarbon; RT-PCR, reverse transcription-polymerase chain reaction; SNP, single nucleotide polymorphism

LUNG CANCER EPIDEMIOLOGY

Globally, lung cancer is the most frequent cancer today and is expected to have a major impact on human health throughout the next decades. The incidence of this epidemic increases by approximately 0.5% each year especially because of contributions from Eastern Europe and developing countries (1,2). In Norway, 1900 new cases of lung cancer were diagnosed in 1997 (3). The role of tobacco smoking as a major etiologic factor of this malignancy is well established (85-90% of lung cancer cases are smokers), and smoking and lung cancer incidence shows a clear dose-response relationship (4,5). However, people may differ in their susceptibility (or resistance) to tobacco smoke carcinogens, since only 15% of life-time smokers get lung cancer. More attention is necessary on the issue of women and smoking because several epidemiological studies have indicated that for a given number of cigarettes smoked, females may be at higher risk of lung cancer compared with males [odds ratio (OR) of 1.2-1.7] (6-8).

GENETIC POLYMORPHISMS AND GENE-ENVIRONMENT INTERACTIONS

There is substantial inter-individual variation in the activity of enzymes that metabolize environmental agents, repair DNA damage, and maintain cell cycle control and immune function. These observations have given rise to the hypothesis that the carcinogenic po-

tential of environmental (and endogenous) agents may be modified by common genetic polymorphisms in the genes that encode these enzymes (9). Interactions between the environment and genetic risk factors (gene-environment interactions) are involved not only in cancer, but in many multifactorial diseases (Figure 1).

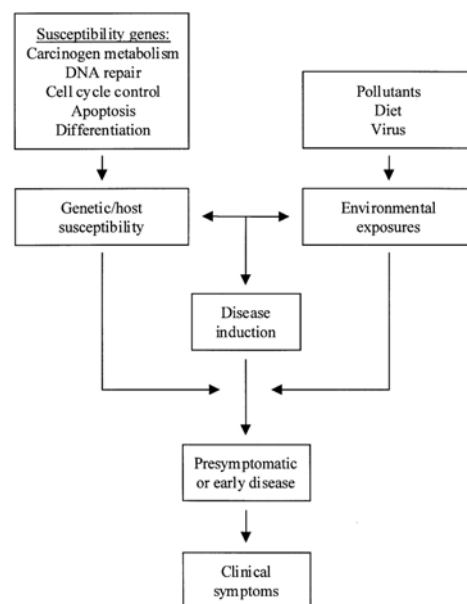


Figure 1. Gene-environment interactions. The risk of adverse effects of environmental exposure may be modified by genetically determined individual susceptibility (reprinted from: 10).

Most human DNA sequence variation is attributable to single nucleotide polymorphisms (SNPs) (11). Genetic polymorphisms are defined as mutations occurring with a frequency of > 1% in the population, but they often have a much higher frequency (> 10%). Such polymorphisms usually have a relatively low penetrance and may be silent in the absence of exposure. However, due to their high frequency, they may be ascribed a large proportion of the particular diseases in the population. This is in contrast to the rare, high penetrance genes (exemplified by Huntington's disease) where the genetic disorder has a strong influence on disease development in the individual, independently of environmental exposure. Somewhere in between these extremes are familial gene disorders, for instance hereditary cancer-predisposition mutations in genes such as *RBI* (retinoblastoma), the *p53* tumor suppressor gene (several cancers), or the adenomatous polyposis coli gene (*APC*, colorectal cancer), where the environment plays a minor role.

Genetic polymorphisms may result in defective protein function and stability, altered posttranscriptional processing, or altered levels of expression. Also, environmental exposures may play a role in determining these parameters, and thereby influence the level of enzyme activity.

CARCINOGEN METABOLISM

Tobacco smoke consists of a complex mixture of more than 3000 compounds of which at least 55 are considered to be carcinogenic (12). Nitrosamines, aromatic amines, and polycyclic aromatic hydrocarbons (PAH)

are important carcinogens in tobacco smoke (13). Most chemical carcinogens are inactive *per se* and require metabolic activation to exert their carcinogenic potential (Figure 2). Carcinogen metabolism is traditionally divided into two distinct phases: The Phase I metabolism, often involving the cytochrome P450 monooxygenase enzymes (CYP), creates reactive electrophiles that may serve as substrates for the second phase. In the Phase II conjugation reactions, less reactive and more water soluble compounds are formed, that may be readily be excreted from the body. These reactions may be carried out by enzymes like the glutathione *S*-transferases (GST) and the *N*-acetyltransferases. Electrophilic intermediates that escape the Phase II detoxification pathway can bind covalently to cellular macromolecules such as proteins and DNA (adducts). If DNA adducts are not repaired by the DNA repair systems, they may be manifested as permanent mutations and may initiate the carcinogenic process.

OUR APPROACH

For several years, we have used smoking-related lung cancer as a model in molecular epidemiological studies of gene-environment interactions. The focus of this review will be on studies we have performed regarding: (1) Individual differences and sex differences in relation to carcinogen metabolism, PAH-DNA adduct levels in the lung, and mutations in the *p53* tumor suppressor gene, and (2) *p53* mutations as molecular biological markers for prognosis in lung cancer patients.

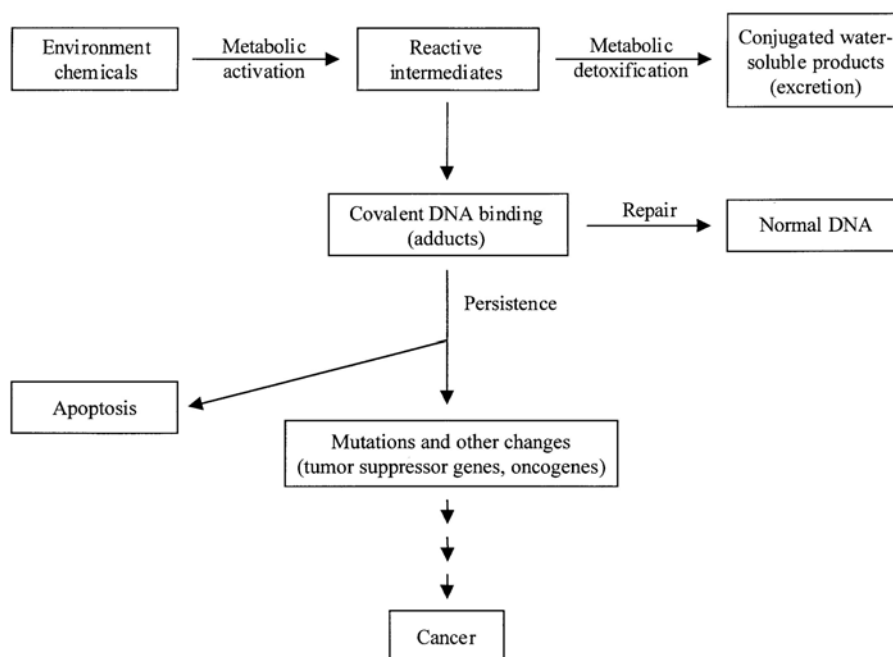


Figure 2. Metabolism of environmental carcinogens, adduct formation, and initiation of carcinogenesis (modified from: 13).

PATIENT MATERIAL

We have, in collaboration with the surgery departments of The Norwegian National Hospital and Haukeland Hospital, collected tissue samples (tumor tissue and peripheral nontumor lung tissue) and blood samples from non-small cell lung carcinoma (NSCLC) patients undergoing surgery. In addition to the histological classification carried out both at the hospitals and in our laboratory, we have from questionnaires obtained data on sex, age, smoking history, occupation, and cancer in the families.

DNA DAMAGE, SMOKING, AND SEX

DNA damage in the form of lung PAH-DNA adducts may be considered a risk factor for the development of lung cancer. In a study of 159 lung cancer cases, PAH-DNA adduct levels were measured in peripheral nontumor lung tissue by the ^{32}P -postlabeling technique. We found that smokers had highly significantly elevated levels of adducts compared to nonsmokers (14). Furthermore, female smokers showed a significantly higher level of adducts (15.4 ± 9.5 adducts/ 10^8 nucleotides, $n = 29$) than male smokers (12.1 ± 8.1 , $n = 93$, $p = 0.047$). Since the females had smoked less than the males, the sex differences became even more significant when expressed as adducts/pack-year (females, 0.95 ± 0.82 ; and males, 0.46 ± 0.46 , $p = 0.0004$) or adducts/cigarette/day (females, 1.48 ± 1.29 ; and males, 0.89 ± 0.74 , $p = 0.015$).

The pulmonary expression of one of the central enzymes in the bioactivation of PAH, CYP1A1, was determined in a subset of the above study population consisting of 15 female and 12 male smokers. We used a quantitative, competitive RT-PCR technique to measure CYP1A1 gene expression (mRNA level), and normalized the data to the expression of the glyceraldehyde 3-phosphate dehydrogenase gene (*GAPDH*) (14). Large inter-individual variations were observed with a total range of 21–1122 CYP1A1 mRNA/ 10^6 GAPDH mRNA. When grouped by sex, we found that female patients had an expression level of 494 ± 334 , which was significantly higher than males (210 ± 201 , Figure 3). Linear regression analysis revealed that the DNA adduct levels were significantly related to the level of CYP1A1 expression. Combined, these data on sex differences in the levels of PAH-DNA adducts and CYP1A1 gene expression lends support to the hypothesis of a higher susceptibility to tobacco carcinogens among females.

MUTATIONAL SPECTRA IN THE *p53* TUMOR SUPPRESSOR GENE

The most common genetic alteration in lung cancer is mutation in the *p53* tumor suppressor gene, which is found in about 60% of the cases (15,16). The *p53* gene is considered an important guardian of the genome,

that coordinates a delicate balance between arrest of the cell cycle to allow repair of DNA damage and apoptosis (programmed cell death or cell suicide) if the damage is severe and irreparable. Mutations are distributed nonrandomly throughout the *p53* gene, and studies have demonstrated a link between certain mutational hot spots in the gene and exposure to specific carcinogens such as PAH (17,18). We have analyzed *p53* mutations in exons 4-9 (within which most mutations occur) in 115 NSCLC tumors using PCR and single-strand conformation polymorphism analysis (19). Mutations were found in 56 of the tumors (49%). A higher frequency of G:C \rightarrow T:A mutations was found in females than in males, a kind of mutation predominantly produced after exposure to PAH (18). This was in spite the fact that the level of exposure to tobacco smoke was lower among the females. Although these findings were not statistically significant, this is in line with other published data (20) and lends further support to the epidemiological evidence of sex differences in lung cancer risk. Our data also showed a significant association between accumulated smoking dose and *p53* mutations.

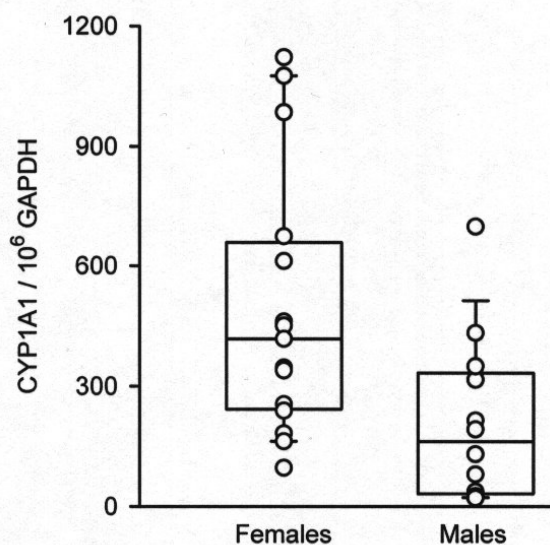


Figure 3. Box plot of lung CYP1A1 gene expression among females ($n = 15$) and males ($n = 12$). Individual CYP1A1 gene expression levels were measured by competitive, quantitative RT-PCR, and normalized to the expression of the GAPDH gene. Solid line within the boxes indicate the median. The box encompasses 25th and 75th percentiles; whiskers extend to 10th and 90th percentiles. Mean \pm SD CYP1A1 expression was 494 ± 334 CYP1A1 mRNA/ 10^6 GAPDH mRNA for females and 210 ± 208 for males ($p = 0.016$, Student's *t* test) (reprinted from: 14).

The prognostic value of the *p53* mutational status was investigated in 148 NSCLC patients as described above (21). Patients with *p53* mutations had a significantly higher risk of dying of lung cancer and death from all causes than those with wild-type *p53* [hazard ratio (HR) = 2.09, 95% confidence interval

(CI) = 1.20–3.64; and HR = 1.69, 95% CI = 1.06–2.70, respectively] (Figure 4). Molecular biological studies have indicated that mutations in codons encoding amino acids that are positioned at specific regulatory domains within the *p53* protein may lead to protein misfunctioning (22). Mutations within a biologically functional group of codons, that have been defined as critical for the conformation of the *p53* protein, were significantly related to shorter lung cancer-related survival (HR = 4.16, 95% CI = 1.93–8.97). These results indicate that mutations within structural and functional domains of *p53* may be useful molecular biological prognostic markers in NSCLC patients (21).

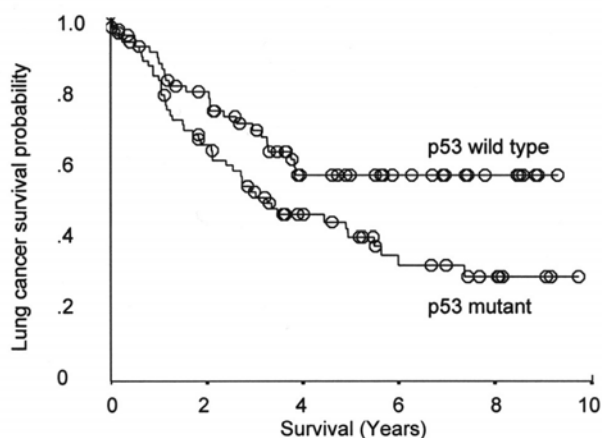


Figure 4. Kaplan-Meier lung cancer related survival curves for NSCLC patients according to *p53* mutational status. Patients with *p53* mutations in exons 4–9 (*p53* mutant, $n = 84$) had significant worse survival than patients with wild type *p53* ($n = 64$, $p = 0.022$) (reprinted from: 21).

GSTM1 AND GSTP1

An important class of Phase II metabolizing enzymes is the glutathione *S*-transferases (GSTs). These enzymes carry out the conjugation of glutathione to electrophilic substances such as reactive intermediates from PAH (reviewed in: 23). *GSTM1* (which is a member of the μ family) is one of the most extensively studied genes concerning metabolic polymorphisms and cancer risk. About 50% of the Caucasian population is homozygous for a deletion in the gene (the *GSTM1* null genotype). The group of individuals with the *GSTM1* null genotype typically has an elevated risk of lung cancer of OR = 1.5 compared to a group carrying one or two intact wild type alleles (24). The major GST in human lung is the *GSTP1* (the π family). A polymorphic site at codon 104 (an A→G substitution that replaces an isoleucine with a valine) in *GSTP1* is known, which has been shown to change the kinetic properties of the enzyme (25).

The codon 104 A→G polymorphism in the *GSTP1* gene was examined in 138 male NSCLC patients and

compared with data from 297 healthy male controls (26). The lung cancer patients were found to have a significantly lower incidence of the AA genotype and a higher incidence of the GG genotype compared to the controls. An OR of 1.90 (95% CI = 1.04–3.47) was obtained when the GG genotype was compared with AG and AA in cases and controls. Lung PAH-DNA adducts were determined in a sub-group of 70 smoking lung cancer patients. The highest DNA adduct level was found in the GG group (15.5 ± 10.2 adducts/ 10^8 nucleotides). However, this was not statistically different from the AG group (12.9 ± 7.4 , $p = 0.39$). Patients with the GG genotype had a significantly higher adduct level than patients with AA (7.9 ± 5.1 , $p = 0.006$). *GSTM1* genotypes were also determined in this study. The *GSTM1* null genotype was associated with a slightly increased risk of lung cancer. In an analysis of combined GST genotypes, it was found that patients with the combination of *GSTM1* null and *GSTP1* AG or GG had significantly higher adduct levels than all other genotype combinations ($p = 0.011$).

The association between *GSTP1* codon 104 genotype and gene expression levels was studied in normal lung tissue from smoking NSCLC patients (S. Mollerup *et al.*, manuscript in preparation). *GSTP1* gene expression was measured by quantitative, competitive RT-PCR, and normalized to the expression of the GAPDH gene. Individual expression levels varied 17-fold, and were apparently not related to sex. The expression level of patients with the AA genotype was significantly higher compared to patients with the GG genotype. Patients with the AG genotype showed an intermediary level of expression. In a recent study of *GSTP1* codon 104 polymorphisms and enzymatic activity in normal lung tissue of lung cancer patients, it was found that patients with the GG genotype had significantly lower activity towards substrates of the *GSTP1* enzyme compared with patients with the AA genotype. Patients with the AG genotype had an intermediary enzyme activity level (27). Together, these data may (at least in part) explain the differences observed in the levels of lung PAH-DNA adducts among the different *GSTP1* codon 104 genotypes. It is, however, not known why this polymorphism, which is positioned within the coding region of the gene, may result in altered gene expression levels.

ADDITIONAL EVIDENCE FOR SEX DIFFERENCES IN LUNG CANCER RISK

There are several lines of molecular biological evidence of a sex difference in risk of lung cancer. In addition to the data obtained in our laboratory, sex differences in the genes or gene products responsible for tobacco smoke carcinogen metabolism (PAH) in humans have been described (28–30). Furthermore, sex differences have been observed in the expression of genes involved in cell growth control (20,31). In a

recent epidemiological study, a higher risk of smoking-associated bladder cancer among females was demonstrated (32). The authors also measured aminobiphenyl-hemoglobin adducts in blood and found that at each level of tobacco smoke exposure, the mean values of these adducts were greater for females than for males. A hypothesis of an involvement of sex steroid hormones in the etiology of lung cancer in females has been forwarded. This has been supported by the observation of an association between lung cancer risk and the use of estrogen replacement therapy in women (33).

Upon entering the cell, PAHs bind to the aryl hydrocarbon receptor (AhR) which then elicits a cascade of intracellular events. Among these events is induction of CYP-enzymes involved in PAH-metabolism. Estrogens exert their cellular effects by binding to estrogen receptors (ER). Activated ERs can then induce the expression of many genes. A complex interaction between AhR and ER signaling pathways has been described. In studies of breast tumor cell lines, it has been demonstrated that the presence of ER may be important for the induction of CYP1A1 via the AhR (34). Little is known about the role of estrogens in lung cancer development. We have analyzed the expression of the two ER types, ER α and ER β , in human lung tissue and tumor cell lines. Both types of receptors appear to be present in human lung, though at different frequency (S. Mollerup *et al.*, manuscript in preparation). We are currently looking further into the possible mechanisms involved in the observed sex differences. In cultures of human lung epithelial cells, we try to manipulate the metabolism of PAH and the expression of relevant genes by various exposures, and by transfecting the ER genes into the cell lines.

FUTURE PERSPECTIVES

There is great interest in developing methods that can give better estimates of (sub)populations or individuals at risk of developing adverse effects after environmental exposures. The metabolism of tobacco smoke carcinogens is complex and involves the action of many genes acting together in pathways. In most cases, polymorphisms in a single gene, or in the expression levels of this gene, only confer a minor variation in the susceptibility of effected groups. The measurement of polymorphisms in one or a few susceptibility genes is therefore of little value in risk assessment. There is a great need for studies addressing overall variations in these susceptibility gene pathways; that is knowledge either about the sum of

genetic polymorphisms within one or more pathways, or on alterations in the levels/activity of the gene products (gene expression).

Recently, the completion of the draft sequence of the human genome was launched (35,36). This event has already imposed a major impact on studies of genetic susceptibility. Not only is the primary sequence of all the genes known, which enables the analysis of their function and regulation, but as a side-effect of the sequencing project, thousands of new genetic polymorphisms have been detected. Of particular interest is to identify SNPs that are of importance for susceptibility.

An especially interesting new technological advancement that has arisen in conjunction with the human genome sequencing project is the microarray technology, which enables the examination of both gene polymorphisms and gene expression profiling (11,37). A cell will respond to a chemical exposure by altered gene expression (and by altered production of proteins) to reduce or neutralize the toxic effect of the insulting agent, and it is assumed that the total sum of gene expression alterations will define the disease-state of the cell.

Toxicogenomics is a new discipline that utilizes genomic data and the emerging array technologies in the study of toxicology. Toxicogenomics may be defined as studies of the effects of toxicants on gene activity and the production of specific proteins by genes in response to these toxicants (38). The toxicogenomics technology may be used in risk assessment to: (1) give better estimates of the toxicity of known chemicals (define their mechanisms of action), (2) predict mechanisms of action and toxicity of new chemicals, and (3) identify subgroups (or individuals) within the population, which are particularly susceptible to environmental chemicals. Furthermore, toxicogenomics may aid in the development of specific drugs for the treatment of diseases, and in the use of the most effective treatment/dose with already available drugs. However, if (when) it becomes possible to identify and measure individual risk parameters with a high level of precision, one must realize that these new technologies also give rise to new and complex ethical and policy issues. Such information will also be of interest for others, for example for employers, insurance companies, and administrations of justice (10).

ACKNOWLEDGEMENTS

The authors wish to thank our collaborators Drs. Lodve Stangeland, David H. Phillips, and Elin Kure for their contributions to this project.

REFERENCES

1. Peto R, Lopez AD, Boreham J, Thun M, Heath C Jr, Doll R. Mortality from smoking worldwide. *Br Med Bull* 1996; **52**: 12-21.
2. Peto R, Chen ZM, Boreham J. Tobacco – the growing epidemic. *Nat Med* 1999; **5**: 15-17.
3. Cancer Registry of Norway. Cancer in Norway: New cases 1997. <http://www.kreftregisteret.no>.
4. IARC. IARC monographs on the evaluation of carcinogenic risk of chemicals to humans, tobacco smoking. *IARC Sci Publ* 1986; **38**: Lyon.
5. IARC. Tobacco. A major international health hazard. *IARC Sci Publ* 1986; **74**: Lyon.
6. Risch HA, Howe GR, Jain M, Burch JD, Holowaty EJ, Miller AB. Are female smokers at higher risk for lung cancer than male smokers? A case-control analysis by histologic type. *Am J Epidemiol* 1993; **138**: 281-293.
7. Engeland A. Trends in the incidence of smoking-associated cancers in Norway, 1954-93. *Int J Cancer* 1996; **68**: 39-46.
8. Zang EA., Wynder EL. Differences in lung cancer risk between men and women: examination of the evidence. *J Natl Cancer Inst* 1996; **88**: 183-192.
9. Rothman N, Wacholder S, Caporaso NE, Garcia-Closas M, Buetow K, Fraumeni JF Jr. The use of common genetic polymorphisms to enhance the epidemiologic study of environmental carcinogens. *Biochim Biophys Acta* 2001; **1471**: C1-C10.
10. Haugen A. Progress and potential of genetic susceptibility to environmental toxicants. *Scand J Work Environ Health* 1999; **25**: 537-540.
11. Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 1998; **280**: 1077-1082.
12. Hoffmann D, Hoffmann I. The changing cigarette, 1950-1995. *J Toxicol Environ Health* 1997; **50**: 307-364.
13. Hecht SS. Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst* 1999; **91**: 1194-1210.
14. Mollerup S, Ryberg D, Hewer A, Phillips DH, Haugen A. Sex differences in lung CYP1A1 expression and DNA adduct levels among lung cancer patients. *Cancer Res* 1999; **59**: 3317-3320.
15. Sidransky D, Hollstein M. Clinical implications of the *p53* gene. *Annu Rev Med* 1996; **47**: 285-301.
16. IARC. The TP53 mutation database. <http://www.iarc.fr>.
17. Greenblatt M, Bennett W, Hollstein M, Harris CC. Mutations in the *p53* tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res* 1994; **54**: 4855-4878.
18. Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science* 1996; **274**: 430-432.
19. Kure EH, Ryberg D, Hewer A, Phillips DH, Skaug V, Baera R, Haugen A. *p53* mutations in lung tumours: relationship to gender and lung DNA adduct levels. *Carcinogenesis* 1996; **17**: 2201-2205.
20. Guinee DG Jr, Travis WD, Trivers GE, De Benedetti VMG, Cawley H, Welsh JA, et al. Gender comparisons in human lung cancer: analysis of *p53* mutations, anti-*p53* antibodies and *C-erbB-2* expression. *Carcinogenesis* 1995; **16**: 993-1002.
21. Skaug V, Ryberg D, Kure EH, Arab MO, Stangeland L, Myking AO, Haugen A. *p53* mutations in defined structural and functional domains are related to poor clinical outcome in non-small cell lung cancer patients. *Clin Cancer Res* 2000; **6**: 1031-1037.
22. Borresen AL, Andersen TI, Eyfjord JE, Cornelis RS, Thorlacius S, Borg A, Johansson U, Theillet C, Scherneck S, Hartman S. TP53 mutations and breast cancer prognosis: particularly poor survival rates for cases with mutations in the zinc-binding domains. *Genes Chromosomes Cancer* 1995; **14**: 71-75.
23. Hayes JD, Pulford DJ. The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit Rev Biochem Mol Biol* 1995; **30**: 445-600.
24. d'Errico A, Taioli E, Chen X, Vineis P. Genetic metabolic polymorphisms and the risk of cancer: a review of the literature. *Biomarkers* 1996; **1**: 149-173.
25. Zimniak P, Nanduri B, Pikula S, Bandorowicz-Pikula J, Singhal SS, Srivastava SK, Awasthi S, Awasthi YC. Naturally occurring human glutathione S-transferase GSTP1-1 isoforms with isoleucine and valine in position 104 differ in enzymatic properties. *Eur J Biochem* 1994; **224**: 893-899.
26. Ryberg D, Skaug V, Hewer A, Phillips DH, Harries LW, Wolf CR, Ogreid D, Ulvik A, Vu P, Haugen A. Genotypes of glutathione transferase M1 and P1 and their significance for lung DNA adduct levels and cancer risk. *Carcinogenesis* 1997; **18**: 1285-1289.
27. 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**: 275-280.
28. Dresler CM, Fratelli C, Babb J, Everley L, Evans AA, Clapper ML. Gender differences in genetic susceptibility for lung cancer. *Lung Cancer* 2000; **30**: 153-160.

29. Tang D, Santella RM, Blackwood AM, Young TL, Mayer J, Jaretzki A, Grantham S, Tsai WY, Perera FP. A molecular epidemiological case-control study of lung cancer. *Cancer Epidemiol Biomarkers Prev* 1995; **4**: 341-346.
30. Tang DL, Rundle A, Warburton D, Santella RM, Tsai WY, Chiamprasert S, Hsu YZ, Perera FP. Associations between both genetic and environmental biomarkers and lung cancer: evidence of a greater risk of lung cancer in women smokers. *Carcinogenesis* 1998; **19**: 1949-1953.
31. Shriver SP, Bourdeau HA, Gubish CT, Tirpak DL, Davis AL, Luketich JD, Siegfried JM. Sex-specific expression of gastrin-releasing peptide receptor: relationship to smoking history and risk of lung cancer. *J Natl Cancer Inst* 2000; **92**: 24-33.
32. Castela JE, Yuan JM, Skipper PL, Tannenbaum SR, Gago-Dominguez M, Crowder JS, Ross RK, Yu MC. Gender- and smoking-related bladder cancer risk. *J Natl Cancer Inst* 2001; **93**: 538-545.
33. Taioli E, Wynder EL. Re: Endocrine factors and adenocarcinoma of the lung in women. *J Natl Cancer Inst* 1994; **86**: 869-870.
34. Thomsen JS, Wang X, Hines RN, Safe S. Restoration of aryl hydrocarbon (Ah) responsiveness in MDA-MB-231 human breast cancer cells by transient expression of the estrogen receptor. *Carcinogenesis* 1994; **15**: 933-937.
35. Craig Venter J, *et al.* The sequence of the human genome. *Science* 2001; **291**: 1304-1351.
36. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001; **409**: 860-921.
37. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995; **270**: 467-470.
38. Nuwaysir EF, Bittner M, Trent J, Barrett JC, Afshari CA. Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 1999; **24**: 153-159.