



Invited article

Scand J Work Environ Health [1999;25\(6\):498-504](#)

doi:10.5271/sjweh.472

Promise of molecular epidemiology - epidemiologic reasoning, biological rationale and risk assessment

by [Vainio H](#)

Key terms: [attributable risk](#); [biomarkers](#); [exposure assessment](#); [gene-environment interaction](#); [somatic mutation](#); [susceptibility marker](#)

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/10884145



This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

Promise of molecular epidemiology — epidemiologic reasoning, biological rationale and risk assessment

by Harri Vainio, MD¹

Vainio H. Promise of molecular epidemiology — epidemiologic reasoning, biological rationale and risk assessment. *Scand J Work Environ Health* 1999;25(6 special issue):498—504.

Molecular epidemiology has emerged as a natural outgrowth of attempts to apply information derived from the explosion in molecular biology to disease in human populations. The incorporation of biomarkers into classical epidemiologic designs holds the promise of unraveling mechanisms, elucidating gene-environment interactions, and dissecting heterogeneity. The primary interest of molecular epidemiology is in the identification of factors in the physical and social environment which affect the risk for disease and which are amenable to preventive intervention. The explosion in molecular technology has not, however, resulted in radical, widespread improvements in epidemiologic results and therefore has led to a sense of frustration in the public health community. As experience accumulates, there is new appreciation that attention to study design, infrastructure, and biomarker validation can improve the results.

Key terms attributable risk, biomarkers, exposure assessment, gene-environment interactions, somatic mutation, susceptibility markers.

The term “molecular epidemiology” refers to the use of biomarkers in epidemiologic studies, with emphasis on markers to measure exposure, characterize host susceptibility, and measure disease. Perera & Weinstein (1) defined “molecular cancer epidemiology” as “advanced laboratory methods in combination with analytical epidemiology to identify at the biochemical or molecular level specific exogenous agents and/or host factors that play a role in human cancer causation”.

Biomarkers are used to make 3 basic types of measurement: (i) internal exposure, often represented by a compound of interest, its metabolite, or a metabolite bonded to a macromolecule (adducts), (ii) early biological effects and disease (“effect markers”), such as mutations and cytogenetic damage which indicate biological effects of pathological significance, and (iii) host susceptibility factors, typically metabolic traits with hereditary

variability in the population, but increasingly focusing on genetic factors.

Molecular assessment of “exposure” in epidemiologic studies

Exposure to chemicals and infectious agents can be measured by the use of molecular biomarkers. Some such measures allow the identification of components of a complex mixture, others confirm that exposure has indeed occurred, and yet others yield an integrated measure of “biologically effective dose” from multiple sources, including internal metabolic processing. In infectious disease epidemiology, the “exposure” can be thought of as contact with an infectious agent. The presence of particular subtypes of the infectious agent can be determined

¹ Unit of Chemoprevention, International Agency for Research on Cancer, Lyon, France.

Reprint requests to: Dr Harri Vainio, Unit of Chemoprevention, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon, cedex 08, France. [E-mail: Vainio@iarc.fr]

by assays of bacterial or viral DNA (deoxyribonucleic acid), allowing more detailed classification of the "exposure".

Biomarkers of biological agents

Biomarkers of exposure extend the reach of classical epidemiology beyond the traditional questionnaire or monitoring methods. Use of biomarkers for infectious agents is, of course, not new in the history of epidemiology; for example, polio antibody patterns were used in early vaccine development to detect immunity. Molecular biomarkers have been used widely in infectious disease epidemiology to differentiate the transmissibility, pathogenicity, and antibiotic sensitivity of, for example, subtypes of the tubercule bacillus, influenza virus, human papillomavirus (HPV), and human immunodeficiency virus. More recently, they have been used in the genetic subtyping of organisms to elucidate their pattern of local and regional spread, including point sources of local clusters of cases. Biomarkers of exposure have been especially useful in studies of the long-term health effects of viruses such as hepatitis B, hepatitis C, and HPV and in studies of *Helicobacter pylori* (2).

A prototype of nucleic-acid-based biomarkers is that for the detection of HPV DNA (2). This marker provides a measure of the presence of type-specific DNA at a given point in time. The test requires the collection of cells with DNA, such as exfoliated cells and biopsy samples, which are analyzed by PCR-based assays (PCR = polymerase chain reaction), considered the method of choice for epidemiologic investigations. One limitation of this marker for cancer epidemiology is that HPV DNA infections are often transient, especially among young women.

Molecular biomarkers have also been used to elucidate causal mechanisms such as that between HPV infection and cervical cancer, which show the interaction between the virus and *p53*. They have also been used to identify new causal agents. For instance, early descriptive epidemiologic data suggested that Kaposi sarcoma (a rare cancer of the skin and connective tissues) affects only a subset of AIDS (acquired immunodeficiency syndrome) patients (ie, homosexual men). Recent analyses of the DNA from Kaposi sarcoma tissue from 27 AIDS patients showed a very high prevalence of herpes viral DNA; healthy tissue samples from the same patients also showed the presence of herpes viral DNA, whereas those of 85 healthy controls did not (3), suggesting that Kaposi sarcoma is due to a previously undescribed herpes virus, which is presumably sexually transmitted.

Macromolecular adducts as biomarkers of exposure to reactive chemicals

Chemicals can bind covalently to cellular macromolecules, such as nucleic acids and proteins. Much attention

has been focused on the measurement of adducts to DNA and proteins, as it has been hypothesized that these not only reflect the relevant exposure but also metabolic activation and the actual quantity of compound that has reached the critical target (4). As the target is DNA, it is plausible that these adducts reflect a biologically relevant event on the pathway to malignancy. Nevertheless, the interpretation of adduct measurements made on human tissues and body fluids must take into consideration the sensitivity and specificity of the measurement, the temporal relationship between exposure and adduct level, and the role of adducts in the process of carcinogenesis. The use of such biomarkers in epidemiologic studies therefore necessitates an optimal study design. The usefulness of adducts is limited, however, by the relatively short half-time of most of those known today.

Biomarkers as measures of early outcomes, predictive of clinical disease

The second possible main use of biomarkers is as measures of early effects that herald the subsequent occurrence of clinical disease. This application is conceptually less straightforward than that for detecting exposure. The "predictivity" of the biomarker must be inferred from knowledge about the natural history of the disease. Biomarkers can potentially be used in 3 main ways: (i) to screen for preclinical disease, (ii) to facilitate conventional epidemiologic studies of disease etiology, and (iii) to monitor variation in health risk.

Cytogenetic changes

Effect markers represent a vast category (5). They include nonspecific markers such as cytogenetically manifest changes in chromosomes (breaks, micronuclei), and more specific findings, such as specific mutations in critical genes like *p53*. The effects can be measured in target tissue such as exfoliated cells from nasal mucosa and urothelium or in more convenient surrogate tissue, such as peripheral white blood cells.

Until recently, the analysis of chromosome alterations relied on conventional staining with DNA-specific stains such as Giemsa. The recent development of in situ hybridization (FISH) techniques has facilitated the detection of specific chromosomes, specific genes, and chromosome alterations. For instance, Lucas et al (6) showed that stable chromosome aberrations could be detected in people decades after exposure to radiation from the atomic bombs in Japan. FISH methods also allow accurate, sensitive assessment of chromosome alterations present in tumors. The specific advance that has made such assessment feasible is known as comparative genomic hybridization (7), which allows the identification of the role

of structural and numerical chromosome alterations in tumor development. The method is being adapted for automated screening approaches involving biochips (8).

Somatic mutations

Few assays are available for detecting gene mutations in humans, whereas somatic cell mutations occur regularly and universally in all people. Some occur "naturally", arising continuously as "spontaneous" replication errors or in response to endogenous mutagens or DNA metabolism, while others are induced by external mutagens that are ubiquitous in the environment. The detection of mutations in the hypoxanthine phosphoribosyltransferase (*HPRT*) gene in lymphocytes is the most extensively employed assay (9). It has been shown in experimental animals that agents that induce cancers in various tissues also produce *hprt* T-cell mutations in vivo. Thus, *hprt* is a functional surrogate for cancer in these species.

At least 7 assays are available for assessing mutations of 5 reporter genes in 2 cell types (10). Mutations scored in red blood cells (ie, in hemoglobin and glycophorin A) occur in nucleated precursor cells, thus limiting the site of mutation in vivo to the bone marrow. Mutations scored in peripheral T cells can arise at any site of the body. As the mutational memory of peripheral T cells is probably a matter of months, at least in adults, mutations in these cells are of no value for detecting remote exposure.

In summary, somatic cell mutations have been shown to occur in vivo in humans, and they are detectable at a variety of loci. The mutants are quantifiable, are defined at the molecular level, and can be detected by highly accurate assays. However, the outstanding issues that remain to be resolved include interindividual variation and the cost and difficulty of the analytical procedures.

Somatic mutations in cancer genes

Malignant transformation is associated with mutations in oncogenes and tumor suppressor genes. Many oncoproteins and tumor suppressor gene proteins are detectable by immunologic techniques like immunoblot and enzyme-linked immunosorbent assay (ELISA) in the body fluids of cancer patients (11, 12). A study of workers exposed in France to vinyl chloride showed that 4 of 5 with liver angiosarcoma and 8 of 9 with liver angiomas had detectable mutant Asp13c-*Ki-ras* p21 (by immunoblotting) in their sera, while none of the 28 unexposed persons had detectable serum mutant p21 protein (13).

The nuclear phosphoprotein encoded by the *p53* tumor suppressor gene is known to accumulate in human tumors in its mutant form. The *p53* gene is mutated in about half of common cancers. Examination of the mutational spectrum of a given target gene should theoretically allow determination of the environmental factors responsible for inducing the mutation. The *p53* gene has proved to be especially suitable for such analysis in that

the analysis of its mutational spectrum has already shown that sunlight has a specific mutagenic role in skin cancer and that aflatoxin is a mutagen involved in liver cancer. About 10% of the mutations in skin tumors are CC>TT base substitutions arising from photodimers at pyrimidine dinucleotides; these are characteristic of damage to DNA caused by ultraviolet radiation. Fewer than 1 per 1000 of internal cancers harbor this type of *p53* mutation (14). Geographically disparate *p53* tumor mutation prevalences and patterns were first shown in relation to hepatocellular cancer in high-incidence regions such as Quidong, China, and in Europe, where mutations are heterogeneous and less frequent (15). A primary risk factor for the hot-spot mutation at codon 249 is the food contaminant aflatoxin, a potent liver carcinogen.

Elevated serum levels of total or mutant *p53* have been reported in patients with hepatocellular carcinoma, breast cancer, lung cancer, or colonic neoplasms. Elevated levels of total or mutant serum *p53* have also been found in asbestosis patients (16) and in vinyl-chloride-exposed workers (17), in persons with and without cancer.

Serum antibodies against *p53* have also been found to occur in patients with different types of cancer and, for instance, in vinyl-chloride-exposed workers. Serum *p53* antibodies were found in a proportion of Finnish lung cancer cases, but most were associated with detectable *p53* mutations in the tumor (18). Current estimates of antibody production in cancer patients according to a simple ELISA range from 5% to 40%. Possible explanations for the selective anti-*p53* response include loss of tolerance due to an accumulation of more stable mutant forms, association of the mutant *p53* with heat-shock proteins, and increased immunogenicity due to conformational tertiary changes induced by specific mutations. As antibody titers are known to drop sharply after therapy, the efficacy of treatment or the presence of an occult recurrence with anti-*p53* titers could be monitored.

Recently, Brauch and his co-workers (19) reported that a unique mutation of von Hippel-Lindau tumor suppressor gene was observed in renal cell cancers of trichloroethylene-exposed people. This is a first "fingerprinting" mutation in renal cell cancer, showing a relationship between exposure to a defined carcinogen (trichloroethylene), specific gene damage, and kidney cancer.

Effect markers as screening tools

Some effect markers might be used to screen for, for example, precursor lesions or early-stage disease (eg, cervical cytology), high-risk persons (eg, prostate-specific antigen), or susceptibility markers (eg, BRCA1). Cervical cytology is well established as an effective means for reducing mortality from invasive cervical cancer. Use of prostate-specific antigen is controversial because it is not specific to invasive prostatic cancer or its precursors; as

it cannot distinguish between precursor lesions that will invade and the majority that will not, a significant amount of overtreatment and morbidity can occur. Furthermore, and crucially, there is no evidence that screening reduces mortality from prostatic cancer.

Susceptibility markers and the identification of high-risk groups

Epidemiologic associations between exposure and disease outcomes have usually been based on the assumption that all persons are equally susceptible to the effects of the exposure, although the effects may not be modified by factors such as age, gender, ethnicity, and hormonal status. Pharmacogenetic studies have shown, however, that people vary in their ability to metabolize drugs (20). There is also increasing evidence that variations in DNA repair capacity, cell cycle control, and immune response may affect the risk for disease. In view of the genetic differences in a number of factors that predict the probability that disease potential will result from an exposure, risk factors may be identifiable only if the association between an exposure and a disease is strong, as in the case of tobacco smoking and lung cancer. Assessments of the risks of populations that have heterogeneous responses may thus be biased and result in risk estimates that are diluted or masked. In studies of the etiology of multicausal diseases, such as cancer, a simplistic approach in which only single factors are evaluated is insufficient, and a multifactorial model is required to evaluate the environmental exposures and genetic and hormonal factors that affect susceptibility. Identification of susceptible subsets of the population, on the basis of the polymorphic genes involved in the line of defense between exposure and the initiation of disease processes in cells may more clearly delineate the factors that increase health risks among some, but not all, persons. New strides may be made in understanding disease etiology and the role of particular factors in etiopathogenesis by conducting molecular epidemiologic studies; however, the incorporation of molecular markers of susceptibility into epidemiologic studies may pose methodological problems that must be addressed by the research community.

Gene-environment interaction

Effect modification, also described as interaction, occurs when the association between an exposure and disease varies with different levels of a 3rd variable. Therefore, data can be stratified on the variable that is thought to modify the effect. Molecular epidemiology has extended the need for stratification to polymorphisms of putative risk-modifying genes. Within specific genetic cate-

gories, associations can be evaluated between groups which are putatively "at risk" and those which are not. This method of studying gene-environment interactions may more clearly elucidate cause and reveal previously unidentified risk factors, by allowing the detection of effects in subgroups when no main effect is observed overall. The data from Ambrosone et al (21) illustrate this concept. Although several studies have shown that tobacco smoking does not seem to increase the risk of breast cancer overall (22), postmenopausal women who have the slow *N*-acetyltransferase 2 (NAT2) genotype and who smoked were at increased risk for breast cancer (21). These findings were not, however, corroborated in the Nurses Health Study (23), which found that cigarette smoking was not appreciably associated with the risk for breast cancer among either slow or fast acetylators.

A similar concept was addressed in 2 recent studies of breast cancer, menopausal status, and the activity of catechol-O-methyltransferase (COMT) (24, 25). The association tended to be null in heterozygous populations, but clear relationships between dependent and independent variables were found when the data were stratified. While Thompson et al (25) found no association between the COMT genotype in a case-control study of pre- and postmenopausal women together, clear, inverse associations were found when the women were stratified by menopausal status. A similar phenomenon was noted by Lavigne et al (24), except that the associations by menopausal status were inverse. Thus the results of the 2 studies are in direct contrast.

A further illustration of the complexities of such studies is represented by investigators of colorectal cancer and the NAT2 genotype. Studies of the role of risk factors such as cigarette smoking have resulted in inconsistent findings. In a study by Welfare et al (26), no association was observed between the NAT2 genotype and cancer risk, but the risk was increased among recent smokers with the slow NAT2 genotype. In contrast, the fast NAT2 genotype was associated with an increased risk among frequent consumers of red meat. This study illustrates the heterogeneity of study populations and the importance of acknowledging gene-environment interactions in studies of cancer risk, so that preventive strategies can be targeted to appropriate groups of people.

Proportion of disease attributable to genotype and the environment

There is the practical question of how knowledge of genetic susceptibility will be used. A danger is that attention will be focused on "defects" in persons rather than on shortcomings in the environment. Is there any preventive advantage of identifying groups of people with a certain genotype (eg, slow acetylators) and removing them from exposure to, for example, benzidine and other aromatic amines? The risk for bladder cancer among sub-

jects who are slow acetylators and are exposed to arylamines has been estimated roughly from the available studies (27). If the relative risk (RR) for slow versus fast acetylators is 2.0 and the proportion of slow acetylators in the exposed population is 50%, the risk attributable (AR) to the slow phenotype is:

$$AR = ((RR-1)/RR) \times Pc,$$

where Pc = the proportion of slow acetylators. Thus,

$$AR = 0.5 \times 0.5 = 0.25 = 25\%.$$

With a relative risk of 2.0, fully 25% of all bladder cancers arising in the exposed population (eg, in the exposed work force) are expected to be attributable to the slow acetylator genotype. If the RR is 10, then the AR is $(10-1)/10 \times 0.5 = 0.45$. Therefore, if all subjects are exposed to the relevant aromatic amines, the AR would vary between 0.0 and 0.5, according to the magnitude of the RR and with a proportion of slow acetylators equal to 0.5.

Removing the slow acetylators from exposure would prevent from 0.0% to 50% of all cancers in the exposed population; however, a more acceptable strategy would be to eliminate the exposure to the carcinogens. In this case the attributable proportion will be simply $(RR-1)/RR$, since everyone is presumed to be exposed to the carcinogen. If the RR of cancer associated with exposure is 2.0, the attributable proportion will be 50%; if it is 10, it will be 90%, and so on. Therefore, for the same magnitude of relative risk — the RR associated with the slow acetylator genotype and the odds ratio associated with exposure to the carcinogens — primary prevention will always be more effective in preventing cancer than genetic screening is. The genetic screening strategy might become more advantageous in arithmetic terms only if the RR for cancer associated with one genotype greatly exceeds the risk associated with the carcinogenic exposure.

Measurement of the predictability of susceptibility markers

There is a heated discussion under way over the use of specific biomarkers aimed to provide predictive tests to identify people considered to be at risk for chronic, non-infectious diseases. Even with the rapid progress in knowledge on susceptibility biomarkers, predictions (in prophetic sense) assume a relevance a posteriori, only after the disease has occurred. Most diseases of public health importance are polygenic disorders, where socio-environmental interactions play a role; their genetic configurations in molecular terms do not therefore allow for a clear identification of predictive biomarkers. Even so, numerous studies have attempted to establish a nexus between attributes involving susceptibility, exposures, and disease, regardless of the contingencies surrounding the predictability phenomena. For the majority of cases,

elements of imprecision make the predictions uninformative, arising from contingencies in dealing with polygenic disorders, variable expression of genetic material, and the unpredictability of the gene-environment interaction. Indicators obtained from most situations in epidemiologic studies consist of mean rates; in the quest for information that can be generalized, one produces an abstract record of individuality, devoid of any reference to a particular person.

Importance of study design and quality control

In order to do informative research, molecular epidemiologists should have both a clear understanding of the significance and methodological pitfalls of the molecular biomarkers under consideration and an appreciation of the importance of the study design. The tail should not be allowed to wag the dog; biomarkers must be used so as to provide better, or faster, answers to epidemiologic research questions. The use of biological markers in epidemiologic research should be a means, not an end (28).

Stratification by a putative effect modifier should provide clearer answers than an overall analysis; however, in many studies of polymorphisms, in, for example, xenobiotic-metabolizing enzymes, and cancer risk, conflicting results have been obtained. The literature of molecular epidemiology is thus filled with inconclusive data, indicating that it is time for the molecular epidemiologic community to explore the areas of bias and flaws in study design and analysis that result in such inconsistent results.

Inadequate power to detect a true effect is one reason for inconsistent results. Statistical power depends on sample size, the size of the effect to be detected, and the variation within the study population. Small sample sizes are common in molecular epidemiologic studies, not only because the assays are expensive and the number of subjects recruited is thus restricted, but also because the method of stratified analysis automatically reduces the population by half or more. For studies in gene-environment interactions, cases and controls are stratified by genotype, and associations between the risk factor and disease status are evaluated separately within each stratum. Even in large studies, therefore, the numbers of subjects in each cell are drastically reduced, increasing the odds for type 1 and type 2 errors. In studies that are necessarily small because of stratification, movement of even a few persons from one category to another could greatly skew the results.

“Molecular epidemiology” should not be viewed as a new, distinct form of scientific enquiry (28). In particular, it should not be considered an alternative orienta-

tion localizing "risk". The primary importance of epidemiologic research, historically and into the foreseeable future, is for identifying factors in the physical and social environment that affect the risk for disease and are amenable to preventive intervention.

The whole human genome will soon have been sequenced, and various research institutes and private laboratories are busy scanning the genome and its polymorphisms. Snippets of DNA from hundreds of persons of different racial backgrounds are sequenced and put into a public repository (29). DNA chips are used to identify polymorphisms in a multitude of genes so that it is possible to evaluate the complex interactions of numerous genetic polymorphisms and environmental exposures. With the advances in the development of tools to identify susceptible subgroups of persons, however, comes the responsibility for devising strategies to ensure validity. This procedure can only be done by considering a number of practical issues at the planning stages of a molecular epidemiology study (30)

References

- Perera FP, Weinstein IB. Molecular epidemiology and carcinogen-DNA adduct detection: new approaches to studies of human cancer causation. *J Chronic Dis* 1982;35(7):581—600.
- Munoz N, Bosch FX. Biomarkers for biological agents. In: Toniolo P, Boffetta P, Shuker DEG, Rothman N, Hulka B, Pearce N, editors. Application of biomarkers in cancer epidemiology. Lyon: International Agency for Research on Cancer (IARC), 1997:127—42. IARC scientific publication, no 142.
- Chang Y, Cesarman E, Pessin MS, Lee F, Culpepper J, Knowles DM, et al. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 1994;266(5192):1865—1969.
- Wild CP, Pisani P. Carcinogen-DNA and carcinogen-protein adducts in molecular epidemiology. In: Toniolo P, Boffetta P, Shuker DEG, Rothman N, Hulka B, Pearce N, editors. Application of biomarkers in cancer epidemiology. Lyon: International Agency for Research on Cancer (IARC), 1997:143—58. IARC scientific publication, no 142.
- Toniolo P, Boffetta P, Shuker DEG, Rothman N, Hulka B, Pearce N, editors. Application of biomarkers in cancer epidemiology. Lyon: International Agency for Research on Cancer (IARC), 1997:1—318. IARC scientific publication, no 142.
- Lucas JN, Awa A, Straume T, Poggensee M, Kodama Y, Nakano M et al. Rapid translocation frequency analysis in humans decades after exposure to ionizing radiation. *Int J Radiat Biol* 1992;62(1):53—63.
- Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray KW, Waldman F, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 1992;258(5083):818—21.
- Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 1997;20(4):399—407.
- Albertini RJ, Nicklas JA, O'Neill Future research directions for evaluating human genetic and cancer risk from environmental exposures. *Environ Health Perspect* 1996;104 suppl 3:503—10.
- Albertini RJ, Hayes RB. Somatic cell mutations in cancer epidemiology. In: Toniolo P, Boffetta P, Shuker DEG, Rothman N, Hulka B, Pearce N, editors. Application of biomarkers in cancer epidemiology. Lyon: International Agency for Research on Cancer (IARC), 1997:159—84. IARC scientific publication, no 142.
- Brandt-Rauf PW. Oncogenes and oncoproteins in occupational carcinogenesis. *Scand J Work Environ Health* 1992;18 suppl 1:27—30.
- Vainio H, Husgafvel-Pursiainen K. Mechanisms of carcinogenesis and molecular epidemiology. *Br J Hosp Med* 1996;56(4):162—70.
- DeVivo I, Marion MJ, Smith SJ, Carney WP, Brandt-Rauf PW. Mutant c-Ki-ras p21 protein in chemical carcinogenesis in humans exposed to vinyl chloride. *Cancer Causes Control* 1994;5(3):273—8.
- Hollstein M, Moeckel G, Hergenhahn M, Spiegelhalter B, Keil M, Werne-Schneider G, et al. On the tumor mutations in cancer genes: insights from the p53 gene. *Mutat Res* 1998;405:145—54.
- Montesano R, Kirby GM. Chemical carcinogens in human liver cancer. In: Brechot C, editor. Primary liver cancer: etiological and progression factors. Boca Raton (FL): CRC Press, 1994:57—77.
- Partanen R, Hemminki K, Koskinen K, Luo JC, Carney WP, Brandt-Rauf PW. The detection of increased amounts of the extracellular domain of the epidermal growth factor receptor in serum during carcinogenesis in asbestosis patients. *J Occup Med* 1994;36(12):1324—8.
- Marion MJ, De Vivo I, Smith S, Brandt-Rauf PW. The molecular epidemiology of occupational carcinogenesis in vinyl chloride exposed workers. *Int Arch Occup Environ Health* 1996;68:394—8.
- Wild CP, Ridanpaa M, Anttila S, Lubin R, Soussi T, Husgafvel-Pursiainen K, et al. p53 antibodies in the sera of lung cancer patients: comparison with p53 mutation in the tumour tissue. *Int J Cancer* 1995;64(3):176—81.
- Brauch H, Weirich G, Hornauer MA, Störkel S, Wöhl T, Bruning T. Trichloroethylene exposure and specific somatic mutations in patients with renal cell carcinoma. *JNCI* 1999;91:854—61.
- Nebert DW. Role of genetics and drug metabolism in human cancer risk. *Mutat Res* 1991;247:267—81.
- Ambrosone CB, Freudenheim JL, Graham S, Marshall JR, Vena JE, Brasure JR, et al. Cigarette smoking, N-acetyltransferase 2 genetic polymorphisms, and breast cancer risk. *JAMA* 1996;276(18):1494—501.
- Baron JA, Newcomb PA, Longnecker MP, Mittendorf R, Storer BE, Clapp RW, et al. Cigarette smoking and breast cancer. *Cancer Epidemiol Biomarkers Prev* 1996;5(5):399—403.
- Hunter DJ, Hankinson SE, Hough H, Gertig DM, Garcia-Closas M, Spiegelman D, et al. A prospective study of NAT2 acetylation genotype, cigarette smoking, and risk of breast cancer. *Carcinogenesis* 1997;18(11):2127—32.
- Lavigne JA, Helzlsouer KJ, Huang HY, Strickland PT, Bell DA, Selmin O, et al. An association between the allele coding for a low activity variant of catechol-O-methyltransferase and the risk for breast cancer. *Cancer Res* 1997;57(24):5493—7.
- Thompson PA, Shields PG, Freudenheim JL, Stone A, Vena JE, Marshall JR, et al. Genetic polymorphisms in catechol-O-

- methyltransferase, menopausal status, and breast cancer risk. *Cancer Res* 1998;58(10):2107—10.
26. Welfare MR, Cooper J, Bassendine MF, Daly AK. Relationship between acetylator status, smoking, and diet and colorectal cancer risk in the north-east of England. *Carcinogenesis* 1997;18(7):1351—4.
 27. Brockmöller J, Cascorbi I, Kerb R, Sachse C, Roots I. Polymorphisms in xenobiotic conjugation and disease predisposition. *Toxicol Lett* 1998;102—103:173—83.
 28. McMichael AJ. Invited commentary—"molecular epidemiology": new pathway or new travelling companion? *Am J Epidemiol* 1994;140(1):1—11.
 29. Marshall E. The human genome project: The hunting of the SNP. *Science* 1997;278:2046—8.
 30. Rothman N, Stewart WF, Schulte PA. Incorporating biomarkers into cancer epidemiology: a matrix of biomarker and study design categories. *Cancer Epidemiol Biomarkers Prev* 1995;4:301—11.