Molecular Epidemiology and Prevention of Cancer

Frederica P. Perera

Columbia University, School of Public Health, New York, New York

Preventable environmental causes of cancer, including tobacco smoke and other carcinogens in the diet, workplace, and ambient environment are responsible for the vast majority of human cancers. This paper reviews recent molecular epidemiologic studies that have focused on environmental carcinogenesis and environment–host interactions. Biomarkers such as carcinogen–DNA and carcinogen–protein adducts, mutations in reporter or target genes (e.g., HPRT, GPA, *ras*, *p53*), or genetic or acquired susceptibility factors (e.g., polymorphisms in the P450 or glutathione-*S*-transferase genes and serum levels of antioxidants) have shown significant potential in prevention. They should be useful in early identification of at risk individuals and in designing and monitoring interventions (smoking cessation, exposure reduction, and chemoprevention). — Environ Health Perspect 103(Suppl 8):233–236 (1995)

Key words: molecular epidemiology, biomarkers, cancer prevention

Introduction

Environmental factors are implicated in the vast majority of human cancers. These include exposures due to lifestyle (smoking, diet, alcohol consumption), and carcinogenic chemicals in the workplace and general environment. In a complex interaction, risk from these environmental exposures is modulated by genetic and acquired susceptibility factors. Genetic factors by themselves are probably responsible for only about 5 to 10% of cancers. Molecular epidemiology has the potential to clarify the contribution of environmental factors to cancer causation and to identify high-risk groups and individuals for purposes of prevention.

A considerable number of carcinogens (agents known to be carcinogenic in experimental systems and/or in humans) are found routinely in the different environmental media. Tobacco smoke contains 4-aminobiphenyl (4-ABP), polycyclic aromatic hydrocarbon (PAH), 4(*N*-methylnitrosamino)-1-(3-pyridil)-1-butanone (NNK), and many other carcinogens. Diverse carcinogens are found in the food supply, including aflatoxin, dichlorodiphenyltrichloroethane (DDT), *N*-nitrosodimethylamine (DMN), PAH, and heterocyclic amines. Table 1 also lists cancer-causing contaminants in the workplace, drinking water, and the ambient air. However, the contaminants listed in Table 1 represent only a small part of the problem because monitoring data are available for a limited number of chemicals. Moreover, an estimated 1000 new chemicals are produced yearly and few of the more than 50,000 chemicals currently in commerce have been sufficiently tested (*1*). From the public health perspective, we need to know what effects these agents are having at the levels at which we are experiencing them, and we need early warning systems to detect those effects.

To prevent environmentally related cancer, we must identify environmental risk factors in a more timely way and then determine which groups and individuals are at greatest risk. The next step is to design interventions targeted to those populations and individuals. Molecular epidemiology has potential in this regard. Molecular epidemiology fuses advances in the molecular biology and molecular genetics of cancer with epidemiology to understand the molecular dose of specific agents, their preclinical biological effects, and the biologic factors that modulate

Table 1. Environmental carcinogens.

Source	Carcinogens identified	Examples
Tobacco smoke	>40	4-ABP, PAH, NNK, EtO, benzene, arsenic
Food	>40	AFB1, DDT/DDE, DMN, PAH, PhIP, EBDC
Workplace	>50	Asbestos, benzene, PAH, arsenic, ETS, 2, 4, 5-T
Drinking water	>40	Benzene, chloroform, alachlor, PAH, TCE, DBCP
Air, indoor/outdoor	>60	ETS, benzene, formaldehyde, PAH, radon, chlordane

Abbreviations: 4-ABP, 4-aminobiphenyl; PAH, polycyclic aromatic hydrocarbon; NNK, 4 (*N*-methylnitrosamino)-1-(3-pyridil)-1-butanone; EtO, ethylene oxide; AFB1, aflatoxin B_1; DDT, dichlorodiphenyltrichloroethane; DDE, dichlorodiphenyldichloroethylene; DMN, *N*-nitrosodimethylamine; PhIP, 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine; EBDC, ethylenebisdithiocarbamates; ETS, environmental tobacco smoke; 2,4,5-T, 2,4,5-trichlorophenoxyacetic acid; TCE, trichloroethylene; DBCP, 1,2-dibromo-3-chloropropane.

This paper was presented at the President's Cancer Panel Conference on Avoidable Causes of Cancer held 7–8 April 1994 in Bethesda, Maryland. Manuscript received 9 March 1995; manuscript accepted 24 March 1995.

This work was supported by U.S. Public Health Service grants 5RO1CA-39174, 5RO1CA-35809, 1-PO1-ESO5294, CA51196, RO1-ES06722, and CA51196; DAMD17-94-J-4251 from the Department of Defense; PDT-373A from the American Cancer Society; and by the Lucille P. Markey Charitable Trust, the Noreen T. Holland Breast Cancer Foundation, the Colette Chuda Environmental Fund, and the March of Dimes Birth Defects Foundation. I thank Gabriela Simon-Cereijido for her assistance in manuscript preparation.

Address correspondence to Dr. Frederica P. Perera, Columbia University, School of Public Health, Environmental Sciences Division, 60 Haven Ave. B-109, New York, NY 10032. Telephone: (212) 305-3465. Fax: (212) 928-2388. E-mail: fpp1@columbia.edu

Abbreviations used: AFB1, aflatoxin B_1; BP, benzo[*a*]pyrene; CA, chromosomal aberrations; CI, confidence interval; DBCP, 1,2-dibromo-3-chloropropane; DDE, dichlorodiphenyldichloroethylene; DDT, dichlorodiphenyltrichloroethane; DMN, *N*-nitrosodimethylamine; EBDC, ethylenebisdithiocarbamates; ELISA, enzyme-linked immunosorbent assay; EtO, ethylene oxide; ETS, environmental tobacco smoke; GPA, glycophorin-A locus; GSTM1, glutathione-*S*-transferase M1; NNK, 4(*N*-methylnitrosamino)-1-(3-pyridil)-1-butanone; OR, odds ratio; PAH, polycyclic aromatic hydrocarbon; PhIP, 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine; SCE, sister chromatid exchange; TCE, trichloroethylene; 2,4,5-T, 2,4,5-trichlorophenoxyacetic acid; 4-ABP, 4-aminobiphenyl; 4-ABP-Hb, 4-aminobiphenyl-hemoglobin.

susceptibility to these exposures (*2,3*). Biomarkers can thereby mitigate the two problems that have plagued epidemiology—the long latency of cancer and fragmentary information on exposure.

Diverse biomarkers have been assayed in human peripheral blood and other tissues. They include internal dosimeters such as cotinine (a metabolite of nicotine) an dichlorodiphenyldichloroethylene (DDE) (a metabolite of DDT), molecular dosimeters such as carcinogen–DNA and carcinogen–protein adducts formed by chemicals such as PAH and 4-ABP, and preclinical effects such as chromosomal aberrations, gene mutations, and changes in the structure or function of oncogenes and tumor-suppressor genes. The extent to which biomarkers are modulated by genetic and/or acquired susceptibility factors such as polymorphisms in the P450 or glutathione transferase genes or serum levels of antioxidant micronutrients have also been studied. Several of our recent studies will be used here to illustrate the advantages and limitations of this approach.

Tobacco smoke is a classic carcinogen and has provided a valuable model for validation of biomarkers, one with direct relevance to cancer prevention. We have recently studied biomarkers in 40 heavy smokers enrolled in a smoking cessation program. They were sampled at enrollment while smoking, 10-weeks later, by which time they had quit, and 6 and 12 months after enrollment. The biomarkers measured in peripheral blood included, cotinine, PAH-DNA adducts in leukocytes, 4-aminobiphenyl-hemoglobin (4-ABP-Hb) adducts, and glycophorin-A (GPA) locus mutation in erythrocytes. Cotinine, PAH-DNA, and 4-ABP-Hb adducts were significantly associated with smoking and declined significantly after individuals quit (Mooney et al., unpublished data). These blood samples were also tested for genetic and acquired factors believed to increase susceptibility—the glutathione-*S*-transferase M1 (*GSTM1*) null genotype and the *CYP1A1*/*Msp*I genotype. Although some epidemiologic studies have concluded that these factors are not associated with increased risk of lung cancer (*4,5*), others have shown a significant association (*6–9*). In subjects with the *CYP1A1*/*Msp*I polymorphism (heterozygotes and homozygotes combined), PAH-DNA adduct levels were only modestly higher than in those with the wild-type *CYP1A1*/*Msp*I. However, a significant association was seen between PAH-DNA adducts and the exon 7 (Ile-Val) genotype. Among the smokers, adduct levels were more than two times higher in subjects with the Ile-Val mutation ($9.8 \pm 8.6/10^8$ nucleotides, $n = 10$) than in subjects without the mutation ($4.5 \pm 5.3/10^8$ nucleotides, $n = 148$) ($p < 0.01$) (*10*). Although limited by small numbers, this study indicates a possible mechanism by which cancer risk from smoking varies within a population.

Passive smoking in the home affects as many as 9 million American children under the age of 5 years (*11*). Passive smoking is associated with respiratory illness and lung cancer in nonsmokers (*12–14*). Moreover, children may be at heightened risk of cancer later in life as a result of exposure to such carcinogens during their early development (*15,16*). Biomarkers can be useful in understanding potential risks of environmental tobacco smoke (ETS) to certain groups of people (including minorities, young children, and women of reproductive age) by providing direct measurements of the internal and biologically effective dose of ETS, its molecular effects, and susceptibility factors that modulate it.

We have recently studied biomarkers among Hispanic and African American preschool children and their mothers with varying exposures to ETS (*17*). Serum cotinine and PAH-albumin adducts were both significantly higher in the children whose mothers smoked than in the children of nonsmoking mothers ($p < 0.001$ and $p < 0.05$, respectively). Cotinine levels were also significantly higher in the children whose mothers did not smoke but who had ETS exposure from other household members than in the unexposed children of nonsmokers. There was a significant dose–response relationship between cotinine and the number of cigarettes smoked per day by the mother both in children ($r^2 = 0.23$, $p = 0.01$) and in the mothers ($r^2 = 0.22$, $p = 0.01$). Among the nonsmoking mothers, regression of biomarkers against total passive smoking exposure also showed a significant association with cotinine ($r^2 = 0.25$, $p = 0.04$). PAH–albumin did not show the same dose-related response with the smoking variables. Mothers' cotinine levels were significantly correlated with those of their children ($r = 0.76$, $p < 0.001$), as were PAH–albumin adducts ($r = 0.27$, $p = 0.014$). These results highlight the need for smoking cessation programs targeted to these populations, particularly to younger women of reproductive age.

A second important environmental exposure is ambient air pollution containing PAH and other products of fossil fuel combustion. In Eastern Europe and other highly industrialized areas, industry is a major source of this type of pollution, as is residential coal burning. Through an international collaborative effort, we have studied biomarkers in residents of the Silesian region in southwest Poland, which is heavily industrialized, has high levels of ambient PAH, and also has high rates of lung and other cancers (*18*). These subjects were compared with residents from a rural area in northeast Poland, which had roughly 10-fold lower levels of air pollution and lower cancer rates. Biomarkers were measured in peripheral blood: PAH-DNA and aromatic-DNA adducts, *ras* oncogene activation, sister chromatid exchange (SCE), and chromosomal aberrations. All were elevated in the environmentally exposed group after adjusting for smoking and age. Further, there were linear relationships among ambient exposures and both measures of DNA adducts as well as SCE and chromosomal aberrations.

So, what do these markers tell us about potential risk? As a first step in answering this question, we conducted a case–control study of non-small cell carcinoma of the lung, with more than 100 cases and a comparable number of controls. A number of biomarkers, including PAH-DNA adducts, were analyzed in peripheral blood from cases and controls (*19*). After adjusting for age, gender, season, ethnicity, and number of cigarettes smoked per day, PAH-DNA adduct levels in white blood cells were significantly associated with lung cancer. The odds ratio (OR) was 2.75 (confidence interval [CI] = 1.37–5.50) when PAH-DNA adduct levels were stratified into high and low categories. Such a finding suggests that the ability to activate and bind carcinogens efficiently may have been a risk factor in the disease.

As mentioned, a polymorphism in the GSTM1 gene has been associated in some but not all (*4,5*) studies with risk of lung cancer (*8,9*). Interestingly, in the case–control study described above, the odds of lung cancer were 12-fold higher in people with both high levels of adducts and the *GSTM1* high-risk genotype ($n = 10$) than in people without these factors ($n = 84$; OR = 11.7, CI = 1.2–115, $p = 0.04$) (*20*). Additional research is needed to confirm these findings.

In a pilot study of biomarkers in breast cancer, DNA adducts were detected in

breast tissue from breast cancer cases and controls by the ^{32}P-postlabeling method using the P1 nuclease extraction procedure (*21*). This method detects aromatic adducts, including those formed by benzo[*a*]pyrene (BP) and other PAHs. Results were available from 31 specimens, including tumor and tumor-adjacent tissue from 15 patients with breast cancer and 4 women undergoing reduction mammoplasty. Aromatic-DNA adduct levels were lower among the noncancer patients than in the cancer patients, although the numbers are too small to draw conclusions. Smoking histories were available on the 15 cases. DNA samples from five of the seven smokers (tumor or tumor-adjacent tissue) displayed the characteristic pattern of smoking-related adducts (a diagonal zone of radioactivity) that has been reported in previous studies of lung cancer patients (*22*). None of the samples from the nonsmokers or from the mammoplasty controls showed this characteristic smoking-related pattern.

These findings are interesting in light of the mixed results from epidemiologic studies of smoking and breast cancer risk. Most studies of active smoking have found either a small positive association or no association with breast cancer (*23*). Some investigators have hypothesized that smoking has the potential to affect breast cancer risk by two competing mechanisms: *a*) reducing risk by altering estrogen metabolism; and *b*) increasing risk by directly exposing breast tissue to genotoxic carcinogens such as PAHs (*24–27*). The data from this pilot study demonstrate that carcinogens found in cigarette smoke, diet, and other environmental media reach the breast and damage DNA. Measurement of carcinogen–DNA adducts as biomarkers of genetic damage may permit identification of women potentially at highest risk from smoking and other carcinogen exposures. Additional samples are now being analyzed.

Biomarkers have the potential to help us understand interindividual variability in cancer risk and differential risks to subsets of the population. An investigation of mothers and newborns in Eastern Europe has shown higher levels of both cotinine and PAH-DNA adducts in the cord blood of newborns than in their mothers, which is quite surprising given that *in utero* exposure to both compounds is likely to be 10-fold lower for the fetus than for the mother. This finding suggests developmentally related susceptibility to tobacco smoke and environmental PAH. The fetus appears to be less protected against DNA damage than the mother, which isconsistent with experimental data.

This study is evaluating gene–environment interactions by measuring the molecular effect of the *CYP1A1* and *GSTM1* genotypes in the newborns. Those infants with the *CYP1A1*/*Msp*I variant genotype ($n = 28$) had higher levels of DNA damage [PAH-DNA adducts by enzyme-linked immunosorbent assay (ELISA)] than those without ($n = 106$). This difference was of borderline significance after controlling for plasma cotinine, place of residence, coal use, diet, and occupational exposures ($p = 0.07$).

Conclusion

Biomarkers have potential in intervention. As shown in the smokenders study, they can document a decline after exposure reduction (*28*). They can also demonstrate modulation of DNA damage by serum antioxidant vitamins. The antioxidants, vitamins C, E, β-carotene, and retinoids, have been shown experimentally to reduce carcinogen activation, adduct formation, mutation, and tumor formation; as some studies indicate a protective effect against human epithelial cancers (*29–31*). An inverse association has also been found between serum levels of several of these micronutrients and carcinogen–DNA adducts (*32*). These results support the usefulness of biomarkers as intermediate end points in monitoring the effects of chemoprevention. In fact, biomarkers are already being incorporated into chemoprevention trials as intermediate end points (*33*).

Although more research is needed to fully validate this approach, the results to date are encouraging and suggest that molecular epidemiology ultimately may prove to be a useful tool in prevention of environmentally related cancer.

REFERENCES

1. Perera FP. Prevention of environmental pollution: good for our health. Environ Health Perspect 101:562–563 (1993).
2. Perera FP. Molecular cancer epidemiology: a new tool in cancer prevention. J Natl Cancer Inst 78:887–898 (1987).
3. Hulka BS, Griffith JD, Wilcosky TC. Biologic Markers in Epidemiology. New York:Oxford University Press, 1990.
4. Nazar-Stewart V, Motulsky AG, Eaton DL, White E, Hornung SK, Leng Z-T, Stapleton P, Weiss N. The glutathione *S*-transferase μ polymorphism as a marker for susceptibility to lung carcinoma. Cancer Res 53:2313–2318 (1993).
5. Shields PG, Bowman ED, Harrington AM, Doan VT, Weston A. Polycyclic aromatic hydrocarbon-DNA adducts in human lung and cancer susceptibility genes. Cancer Res 53:3486–3492 (1993).
6. Kihara M, Kihara M, Kazumasa N. Lung cancer risk of *GSTM1* null genotype is dependent on the extent of tobacco smoke exposure. Carcinogenesis 15:415–418 (1994).
7. Anttila S, Hirvonen A, Husgafvel-Pursiainen K, Karjalainen A, Nurminen T, Vainio H. Combined effect of CYP1A1 inducibility and *GSTM1* polymorphism on histological type of lung cancer. Carcinogenesis 15:1133–1135 (1994).
8. Brockmoller J, Kerb R, Drakoulis N, Nitz M, Roots I. Genotype and phenotype of glutathione *S*-transferase class μ isoenzymes μ and ψ in lung cancer patients and controls. Cancer Res 53:1004–1011 (1993).
9. Zhong S, Howie AF, Ketterer B, Taylor J, Hayes JD, Beckett GJ, Wathen CG, Wolf CR, Spurr NK. Glutathione *S*-transferase μ locus: use of genotyping and phenotyping assays to assess association with lung cancer susceptibility. Carcinogenesis 9:1533–1537 (1991).
10. Mooney LA, Bell DA, Lucier G, Ottman R, Santella RM, Tsai WY, Covet L, Young TL, Perera FP. Modulation of DNA adducts in heavy smokers by genetic factors. Proc Am Assoc Cancer 36:120 (1995).
11. ACS. Cancer Facts and Figures—1993. Atlanta:American Cancer Society, 1993.
12. DHHS. The Health Consequences of Involuntary Smoking. A Report of the Surgeon General. Washington:U.S. Department of Health and Human Services, 1986.
13. National Research Council. Environmental Tobacco Smoke: Measuring Exposures and Assessing Health Effects. Washington:National Academy Press, 1986.
14. U.S. EPA. Respiratory health effects of passive smoking: lung cancer and other disorders. Rpt No EPA/600/6-90/006F. Washington:U.S. Environmental Protection Agency,1992.
15. National Research Council. Pesticides in the Diets of Infants and Children. Washington:National Academy Press, 1993.
16. Janerich DT, Thompson WD, Varela LR, Greenwald P, Chorost S, Tucci C, Zaman MB, Melamed MR, Kiely M, McKneally MF. Lung cancer and exposure to tobacco smoke in

the household. N Engl J Med 323:632–636 (1990).
17. Crawford FG, Mayer J, Santella RM, Cooper T, Ottman R, Tsai WY, Simon-Cereijido G, Wang M, Tang D, Perera FP. Biomarkers of environmental tobacco smoke in preschool children and their mothers. J Natl Cancer Inst 86:1398–1402 (1994).
18. Perera FP, Hemminki K, Grzybowska E, Motykiewicz G, Michalska J, Santella RM, Young TL, Dickey C, Brandt-Rauf P, DeVivo I, Blaner W, Tsai W-Y, Chorazy M. Molecular and genetic damage from environmental pollution in Poland. Nature 360:256–258 (1992A).
19. Tang D, Santella RM, Blackwood A, Young TL, Mayer J, Jaretzki A, Grantham S, Carberry D, Steinglass KM, Tsai WY, et al. A case-control molecular epidemiologic study of lung cancer. Cancer Epidemiol Biomarkers Prev 4:341–346 (1995).
20. Tang D, Shiamprasert S, Santella RM, Young TL, Tsai WY, Perera FP. Molecular epidemiologic risk model for lung cancer. Proc Am Assoc Cancer Res 36:284 (1995).
21. Perera FP, Estabrook A, Hewer A, Channing KM, Rundle A, Mooney LA, Whyatt R, Phillips DH. DNA damage from environmental carcinogens in human breast tissue. Cancer Epidemiol Biomarkers Prev 4:233–238 (1995).
22. Phillips DH, Hemminki K, Alhonen A, Hewer A, Grover PL. Monitoring occupational exposure to carcinogens: detection by ^{32}P-postlabeling of aromatic DNA adducts in white blood cell DNA from foundry workers. Mutat Res 204:531–541 (1988).
23. Palmer JR, Rosenberg L. Cigarette smoking and the risk of breast cancer. Epidemiol Rev 15:145–156 (1993).
24. Hiatt RA, Fireman BH. Smoking, menopause, and breast cancer. J Natl Cancer Inst 76:833–838 (1986).
25. ACS. Annual Report. 1988 Cancer Facts and Figures. New York:American Cancer Society, 1989.
26. Chu SY, Stroup NE, Wingo PA, Lee NC, Peterson HB, Gwinn ML. Cigarette smoking and the risk of breast cancer. Am J Epidemiol 131:244–253 (1990).
27. Palmer JR, Rosenberg L, Clarke EA, Stolley PD, Warshauer ME, Zauber AG, Shapiro S. Breast cancer and cigarette smoking: a hypothesis. Am J Epidemiol 134:1–13 (1991).
28. Mooney LA, Santella RM, Covey L, Jeffrey AM, Bigbee W, Cooper TB, Ottman R, Tsai WY, Wazneh L, Glassman AH, Young TL, Perera FP. Decline in DNA damage and other biomarkers in peripheral blood following smoking cessation. Cancer Epidemiol Biomarkers Prev (in press).
29. Perera FP, Tang D, Grinberg-Funes RA, Blackwood A, Dickey C, Blaner W, Santella RM. Molecular epidemiology of lung cancer and the modulation of markers of chronic carcinogen exposure by chemopreventive agents. J Cell Biochem (Suppl) 17F:119–128 (1993).
30. Block G. The data support a role for antioxidants in reducing cancer risk. Nutr Rev 50:207–213 (1992).
31. Willett WC. Vitamin A and lung cancer. Nutr Rev 48:201–211 (1990).
32. Grinberg Funes RA, Singh VN, Perera FP, Bell DA, Young TL, Dickey C, Wang LW, Santella RM. Polycyclic aromatic hydrocarbon-DNA adducts in smokers and their relationship to micronutrient levels and glutathione-*S*-transferase M1 genotype. Carcinogenesis 15:2449–2454 (1994).
33. Kelloff GJ, Malone WF, Boone CW, Steele VE, Doody LA. Intermediate biomarkers of precancer and their application in chemoprevention. J Cell Biochem (Suppl) 16G:15–21 (1992).