Journal of Cellular Biochemistry 25S:63–68 (1996)

Application of Molecular Epidemiology to Lung Cancer Chemoprevention

LaVerne A. Mooney* and Frederica P. Perera*

Columbia University School of Public Health, Division of Environmental Health Sciences, New York, New York

Abstract Molecular epidemiology has made great progress in detecting and documenting carcinogenic exposures and host susceptibility factors, in an effort to explain interindividual variation in disease. Interindividual differences in cancer risk have been hypothesized to result from an array of both genetic and acquired factors including nutritional status. Elevated risk of lung cancer has been associated with polymorphisms of metabolic genes such as CYP1A1 and GSTM1. On the other hand, numerous studies have demonstrated that diets rich in fruits and vegetables are protective against cancer, and have correlated high levels of antioxidants in the blood with decreased risk.

As a first step in identifying susceptible individuals, we have assessed the combined effect of genetic factors and nutritional status on DNA adducts in a population of healthy smokers. Plasma retinol, β-carotene, α-tocopherol, and zeaxanthin were inversely correlated with DNA damage, especially in subjects lacking the "protective" GSTM1 gene. Research is ongoing using biomarkers to determine the effect of supplementation with antioxidants/vitamins on DNA damage, especially in population subsets with putative "at risk" genotypes. Information on mechanisms of interactions between exposure, micronutrients, and other susceptibility factors is important in the development of effective practical interventions. J. Cell. Biochem. 25S:63–68.

Key words: antioxidants; chemoprevention; DNA damage; genotype; lung cancer

Molecular epidemiology is a discipline that merges highly sophisticated laboratory techniques with epidemiologic methods, producing a powerful tool that can be used in cancer prevention [1]. In the past 10 years, molecular epidemiology has made substantial progress in validating biomarkers in populations with well-defined exposures and/or with a defined risk of cancer. Molecular markers or biomarkers have been used to identify hazards and assess dose-response relationships in populations with exposure to carcinogens via ambient air pollution, occupation, and/or lifestyle (diet, smoking, etc.) [2,3]. In addition, molecular markers of susceptibility and early biologic response have the potential to identify high-risk populations as candidates for interventions such as exposure reduction and chemoprevention. Examples of validated biomarkers include: internal and molecular dosimeters of carcinogens (e.g., carcinogen-DNA adducts), alterations in oncogenes and tumor suppressor genes (e.g., *ras* and p53) and genetic factors such as polymorphisms in specific genes (e.g., CYP1A1, GSTM1) involved in the metabolism of carcinogens.

A major goal of molecular epidemiology is to elucidate the mechanisms that explain why persons with the same apparent exposure vary in their risk of disease. Interindividual variation in response is thought to be due to interactions between environmental exposures and host susceptibility factors. Once these interactions are identified, mechanistically relevant biomarkers could be used as "targets" or monitors of interventions in susceptible populations [4–9].

Using lung cancer as a model, this paper will illustrate the potential benefits of applying mo-

Contract grant sponsor: NIEHS, contract grant number 1-P01-ES05294; Contract grant sponsor: NIDA, contract grant number NIDA5-R-DA04732; Contract grant sponsor: American Cancer Society, contract grant number Sig–13, contract grant number PDT-373; Contract grant sponsor: Cancer center core grant; Contract grant sponsor: NIH, contract grant number NIH CA51196, contract grant number NIH 5 PO1 CA53772; Contract grant sponsor: Irving A. Hansen Memorial Foundation; Contract grant sponsor: Lucille P. Markey Charitable Trust.

*Correspondence to: Drs. Frederica P. Perera and LaVerne A. Mooney, Columbia University School of Public Health, Division of Environmental Health Sciences, 60 Haven Avenue B-1, New York, NY 10032.

Received 18 October 1995; Accepted 15 March 1996

lecular epidemiology to chemoprevention. A brief summary of environmental and host susceptibility risk factors for lung cancer will be presented, followed by a discussion of several recent molecular epidemiologic studies in high-risk populations. Criteria for using biomarkers in chemoprevention and future studies will be described.

ENVIRONMENTAL RISK FACTORS FOR LUNG CANCER

Cigarette smoke has been identified as the major environmental cause of lung cancer [10]. Epidemiologic studies now estimate that smoking is responsible for > 80% of all lung cancers [11]. Exposure to radon, asbestos and organic solvents have also been associated with increased risk of lung cancer, alone or in combination with smoking [10,11]. Although the exposures associated with the disease are well-documented, there is still tremendous interindividual variability in lung cancer risk in smokers. The odds of getting lung cancer for smokers are more than 10 times higher than those for nonsmokers, yet only a fraction of smokers will be diagnosed with the disease in their lifetime [12]. In recent years the ability to analyze human samples for genetic polymorphisms, DNA damage, nutrient status, and oncoproteins has shed light on the sources of variability that may indicate inherited or acquired susceptibility. Although many of the exposure risk factors were individually identified for lung cancer, the state of knowledge has advanced to the point where these environmental risk factors are being investigated in conjunction with multiple host susceptibility factors so that gene-environment interactions can be assessed.

KNOWN OR SUSPECTED LUNG CANCER SUSCEPTIBILITY FACTORS

Genetic factors play an important role in influencing individual susceptibility to a variety of cancers. Susceptibility to disease may result from modulation of metabolic pathways that activate and detoxify carcinogens, inactivation of tumor suppressor genes, poor DNA repair capacity or reduced immunologic function, and nutrient status [4,5,13]. Genetic predisposition to cancer induction may also result from inherited mutations in tumor suppressor genes (e.g., retinoblastoma or p53) which regulate cell growth and terminal differentiation [for review see 14,15].

While most of these susceptibility variables are substantially under genetic control, environmental or acquired factors may also play a role. For example, many carcinogens are metabolically activated or detoxified by P450 enzymes before binding to DNA. Thus, a number of studies indicate that elevated metabolic activity of a specific cytochrome P450 enzyme (CYP1A1) is associated with lung cancer risk [reviewed in 16]. Moreover, although studies in Caucasians have not seen the same strong effect, two closely-linked polymorphisms in the CYP1A1 gene (exon 7 variant polymorphism and MspI RFLP) have been associated with a 2–3-fold increased risk of lung cancer in Japanese and Brazilian homozygous individuals [17–19]. Japanese lung cancer patients with the MspI variant genotype have also reported a lower lifetime consumption of cigarettes than lung cancer patients without the genotype, suggesting increased susceptibility at lower doses [18]. The exon 7 variant protein (valine) has been correlated with increased CYP1A1 enzyme activity *in vitro* supporting the biological plausibility of this polymorphism in cancer [20].

Lack of activity of the "phase II" detoxifying enzyme glutathione-S-transferase M_1 (GSTM1) has been associated with a 2–3 fold greater risk of developing lung adenocarcinoma in some but not all, case-control studies [reviewed in 21]. A recent metanalysis of these studies yielded an average increased risk of lung cancer of 1.4 for subjects with the GSTM1 deletion [21]. In Japanese subjects, the combination of both CYP1A1 exon 7 and GSTM1 at risk genotypes resulted in a greater than additive increased risk for all histological types of lung cancer 5.8 (95% CI = 2.3–13.3), and a greater than multiplicative relative risk of 9.1 (95% CI = 3.4–24.4) for squamous cell carcinoma [18]. These results highlight the importance of assessing multiple polymorphisms.

Another factor which may influence susceptibility to certain cancers is an individual's nutritional status, resulting from intake of dietary fat and vitamins [22–24]. Epidemiologic studies are convincing of a protective effect of fruits and vegetables rich in vitamins C, E, and β-carotene for epithelial cancers in many organs including the lung [reviewed in 24]. Twenty-nine of 31 studies showed a significant protective effect for vitamin C and β-carotene and lung

cancer [24]. Overall, for the major epithelial cancers, 120/130 studies showed a statistically significant reduction in risk by vitamins C, E, β-carotene or food sources rich in these micronutrients. By contrast, the epidemiologic evidence for a protective effect of retinol is less compelling [23,24]. More recently, studies indicate that treatment with high doses of β-carotene may not be protective in heavy smokers and may even be harmful (NI in press).

The experimental evidence for a protective effect of micronutrients in lung cancer is reviewed elsewhere [4,24]. The available experimental data largely concern the antioxidants and free radical scavengers (vitamin C, β-carotene, and the carotenoids) and retinol. Micronutrients, including the antioxidants mentioned above and retinol, are involved in a variety of functions and mechanisms that affect metabolism of endogenous and exogenous chemicals. They may act directly to quench oxidants, including polycyclic aromatic hydrocarbons (PAHs), and thereby reduce DNA damage. In addition, they may operate indirectly by modulating immune function or gene expression. There is also evidence that retinol and the retinoids enhance cell differentiation, suppress malignant transformation and counteract the effect of tumor promoters. Experimentally, they reduce the production of DNA adducts, DNA damage, mutation and/or SCE by diverse carcinogens (including benzo(a)pyrene (BP), aflatoxin B_1 (AFB_1), and N-nitrosodimethylamine (DMN)). In human lung cancer cells, retinoids increased the expression of the tumor suppressor gene, p53.

Both antioxidants and retinoids have been reported to impact the expression of metabolic enzyme activity. For example, antioxidants increase P450 levels in certain tissues. Activity of the cytochrome P450-dependent mixed-function oxidases, the major enzyme system involved in BP metabolism and activation, has been found to decrease with vitamin A deficiency. Antioxidants induce glutathione-S-transferase (GST) in vitro; and GST activity is reportedly diminished in various cells and tissues with vitamin A deficiency.

In humans there appear to be redundant mechanisms to evade DNA damage. A study of smokers suggests that α-tocopherol may be the primary antioxidant defense in the lungs in response to cigarette smoke exposure [25]. In plasma, vitamin C acts directly as an antioxidant in cigarette smoke-exposed subjects [26]. In some but not all studies, vitamins and/or GSTM1 appear to protect against DNA damage [27–29].

BIOMARKERS IN CHEMOPREVENTION

Biologic markers already play an important role in the evaluation of chemopreventive agents, specifically in Phase II trials [8,30,31]. The most commonly used genetic markers in current trials involving lung and upper aerodigestive tract tumors include micronuclei, DNA content, and genetic alteration in oncogenes, as well as markers of proliferation, growth regulation and differentiation [31]. Research suggests that additional biomarkers of genetic damage such as carcinogen-DNA adducts can be useful in intervention studies of exposed, or "at risk" populations.

DNA and protein-carcinogen adducts are molecular markers that appear early in carcinogenesis. Adducts can be useful as they reflect not only exposure, but host genetic and nutritional susceptibility and metabolic capacity as well. They are frequent events that occur early enough in the disease process to provide the opportunity for intervention. Compared to DNA adducts, oncogene activation and tumor suppressor gene mutations tend to occur later in the disease process and hence are less common in subjects without apparent clinical disease. On the other hand, susceptibility factors such as antioxidants and genetic polymorphisms may act early or late in the process of cancer development.

Benzo(a)pyrene (BP) is an environmentally ubiquitous exogenous carcinogen found in cigarette smoke and ambient air. BP is metabolically activated by the P450 (mainly CYP1A1) enzymes and detoxified by GSTM1 enzymes. In certain instances, BP intermediate metabolites are formed that can bind covalently to DNA, sometimes resulting in characteristic mutations. Studies of BP-induced rodent lung tumors and in vitro mammalian cell assays have demonstrated the ability of benzo(a)pyrene diol epoxide to predominantly cause G→t transversions [5]. Transversions (G→t) in the p53 gene are the most common base substitution observed in lung cancer cases [32]. This distinct mutation was found in approximately half of all non-small cell lung cancers, and was positively associated with lifetime cigarette consumption [33]. The G→t mutations are also found in the

K-*ras* protooncogene of human lung adenocarcinomas. The ability of BP to bind to DNA and preferentially cause this type of mutation in critical genes is supportive of the hypothesis that smoking-related mutations could be instrumental in lung carcinogenesis by activating oncogenes and inactivating tumor suppressor genes. Although such data are not conclusive, they are useful in generating hypotheses with respect to mechanisms of environmental carcinogenesis.

RECENT MOLECULAR EPIDEMIOLOGIC STUDIES IN SUBJECTS AT HIGH RISK OF LUNG CANCER

Two recent molecular epidemiologic studies have examined the relationship between levels of micronutrients (vitamins C, E, β-carotene, retinol, and the carotenoids) in peripheral blood on the one hand, and biomarkers of genetic damage on the other, in populations of smokers who were not supplemented with vitamins. The measure of DNA damage assessed (PAH-DNA adducts) was previously validated as a measure of exposure to environmental carcinogens/mutagens in cigarette smoke.

In a cross-sectional study of 63 healthy current smokers, plasma concentrations of α-tocopherol and vitamin C were significantly inversely correlated with PAH-DNA adducts in mononuclear leukocytes measured in the same individuals [29]. The inverse association between adducts and α-tocopherol was stronger in those with the GSTM1 null genotype than in those with the gene present [29]. Plasma vitamin C levels were also inversely correlated with DNA adducts; the results were of marginal statistical significance among subjects with the GSTM1 null genotype and not significant in the others.

In a second cross-sectional study of 159 heavy smokers entering a smoking cessation program, PAH-DNA adducts in total white blood cells (drawn while subjects were still smoking) were significantly higher (2-fold) in subjects with the exon 7 (Ile-Val) variant genotype compared with those without, but were only slightly elevated in subjects who had the MspI RFLP polymorphism but without the exon 7 variant (in press). These results, in conjunction with the in vitro and lung-cancer studies mentioned above support the hypothesis that the exon 7 polymorphism is the functional mutation [19,20]. In the same subjects, plasma levels of retinol, α-tocopherol, and β-carotene were inversely associated with PAH-DNA adduct levels in total white blood cells. PAH-DNA adducts were inversely and significantly associated with retinol levels in all subjects, and with smoking-adjusted β-carotene levels in the subset of subjects with the GSTM1 null genotype. Similarly, in serial samples from a subset of 40 subjects who were able to quit smoking, a significant inverse relationship between adducts and vitamins (α-tocopherol and zeaxanthin) was observed in GSTM1 null subjects (in press). These results suggest that some individuals may be at increased risk of DNA damage due to a combination of low plasma antioxidant/micronutrient levels and susceptible genotype.

CRITERIA FOR BIOMARKERS IN CHEMOPREVENTION STUDIES

The criteria for the inclusion of biomarkers in chemoprevention studies have been described by others: biomarkers should be measurable and quantitative, have a high positive predictive value (PPV), and be modulated by a chemopreventive agent [3,8,9]. These criteria have been adapted to include markers that have been validated as early response markers. Due to their placement in the scheme of carcinogenesis, markers such as adducts may not have a high positive predictive value (PPV), but can reflect individual response to exposure. If they are also clearly modulated by antioxidants/vitamins they have potential in chemoprevention studies, preferably coupled with preclinical effect markers such as altered oncogenes and tumor suppressor genes or cell proliferation and differentiation.

CONCLUSIONS AND FUTURE STUDIES

Molecular mechanisms of susceptibility or protection can be hypothesized from *in vitro* and epidemiologic studies, and promising hypotheses can be tested in a molecular epidemiologic framework. For lung cancer, one mechanism appears to involve both antioxidants and susceptibility genes. The use of validated markers of DNA damage (e.g., DNA adducts) or biological effect markers (e.g., oncoproteins), in combination with genetic and nutritional markers can help to confirm or refute mechanistic hypotheses. Molecular markers are especially useful when there is effect modification (e.g., where only those with susceptible genotypes are affected and there is little or no effect in the

population as a whole). Although markers such as DNA damage may not be strongly related to risk, they may be useful in generating information on the metabolic response to carcinogenic exposure and chemopreventive agents.

In addition, biomarkers can be useful in pilot studies to determine their modulation by a chemoprevention agent, the kinetics of the response to the chemopreventive agent, and the duration of the effect. Since the efficiency of interventions may be increased by studying high risk groups, markers of susceptibility may be used to identify these populations for initial studies.

Biological effect markers should also be validated as early response markers or risk markers. This may best be accomplished by a nested-case control design within an ongoing lung cancer chemoprevention trial where cohorts have been optimally sampled for the biomarkers of interest. Markers with high positive predictive value may then be used instead of, or in addition to, the cancer endpoint. Given the complexity of carcinogenesis, a combination of biomarkers will improve our understanding of the disease process and our ability to monitor chemoprevention trial efficacy. Use of several markers will increase the likelihood that the biomarkers will reflect the mechanisms of the chemoprevention agent.

Finally, wherever possible, biomarkers should be incorporated into ongoing large-scale chemoprevention trials, when sampling protocol and design considerations allow. The use of such markers may ultimately reduce sample size and time required to test promising chemopreventive agents.

Molecular epidemiology has achieved several of its stated goals. It has been useful in identifying exposures and high risk groups, and in understanding molecular mechanisms of biologic response. Now, it is poised for application to the field of cancer chemoprevention.

ACKNOWLEDGMENTS

This work was supported by grants from NIEHS 1-PO1-ESO5294, NIDA-5-R-DA04732, the American Cancer Society Sig-13 and PDT-373, Cancer center core grant, NIH CA51196, NIH 5 PO1 CA53772, the Irving A. Hansen Memorial Foundation, and an award from Lucille P. Markey Charitable Trust.

We thank Gabriela Simon-Cereijido for her assistance in manuscript preparation.

REFERENCES

1. Schulte PA, Perera FP (eds) (1993): "Molecular Epidemiology: Principles and Practices." New York: Academic Press.
2. Perera FP: Molecular cancer epidemiology (1987): A new tool in cancer prevention. J Natl Cancer Inst 78:887–898.
3. Hulka BS (1991): Epidemiological studies using biological markers: Issues for epidemiologists. Cancer Epidemiol Biomarkers Prev 1:13–19.
4. Perera FP, Mooney LA (1993): The role of molecular epidemiology in cancer prevention. In DeVita VT, Hellman S, Rosenberg SA (eds): "Cancer Prevention." Philadelphia: J.B. Lippincott, pp 1–15.
5. Harris C (1991): Chemical and physical carcinogenesis: Advances and perspectives for the 1990s. Cancer Res 51 (Suppl):5023S–5044S.
6. Guidotti TL (1992): Hidden assumptions in environmental research. Environ Res 59:101–113.
7. Benner SE, Pastorino U, Lippman SM, Hong WK (1994): Second international cancer chemoprevention conference. Cancer Res 54:854–856.
8. Kelloff GJ, Malone WF, Boone CW, Steele VE, Doody LA (1992): Intermediate biomarkers of precancer and their application in chemoprevention. J Cell Biochem 16G (Suppl):15–21.
9. Greenwald P, Nixon DW, Malone WF, Kelloff GJ, Stern HR, Witkin KM (1990): Concepts in cancer chemoprevention research. Cancer 65:1483–1490.
10. Doll R, Peto R (1981): "The Causes of Cancer: Quantitative Estimates of Avoidable Risk of Cancer in the United States Today." New York: Oxford University Press, pp 1220–1224.
11. ACS (1994): "Cancer Facts and Figures—1994." Atlanta: American Cancer Society.
12. Villeneuve PJ, Mao Y (1994): Lifetime probability of developing lung cancer, by smoking status Canada. Can J Public Health 85:385–388.
13. Vineis P, Caporaso N (1995): Tobacco and cancer: Epidemiology and the laboratory. Environ Health Perspect 103:156–160.
14. Marshall CJ (1991): Tumor suppressor genes. Cell 64: 313–326.
15. Li FP (1990): Familial cancer syndromes and clusters. Curr Probl Cancer 49:75–113.
16. Vahakangas K, Pelkonen O (1989): Host variations in carcinogen metabolism and DNA repair. In Lynch HT, Hirayama T (eds): "Genetic Epidemiology of Cancer." Boca Raton: CRC Press, pp 35–54.
17. Hayashi S, Watanabe J, Kawajiri K (1992): High susceptibility to lung cancer analyzed in terms of combined genotype of P450IA1 and *Mu*-class glutathione *S*-transferase genes. Jpn J Cancer Res 83:866–870.
18. Nakachi K, Imai K, Hayashi S, Watanabe J, Kawajiri K (1991): Genetic susceptibility to squamous cell carcinoma of the lung in relation to cigarette smoking dose. Cancer Res 51:5177–5180.
19. Hamada GS, Sugimura H, Suzuki I, Nagura K, Kiyokawa E, Iwase T, Tanaka M, Takahashi T, Watanabe S, Kino I, Tsugane S (1995): The heme-binding region polymorphism of cytochrome P450IA1 (*CypIA1*), rather than the *Rsa*I polymorphism of IIE1 (*CypIIE1*), is associated with lung cancer in Rio de Janeiro. Cancer Epidemiol Biomarkers Prev 4:63–67.

20. Crofts F, Taioli E, Trachman J, Cosma GN, Currie D, Toniolo P, Garte SJ (1994): Functional significance of different human CYP1A1 genotypes. Carcinogenesis 15:2961–2963.
21. McWilliams JE, Sanderson BJS, Harris EL, Richert-Boe KE, Henner WD (1995): Glutathione *S*-transferase M1 (GSTM1) deficiency and lung cancer risk. Cancer Epidemiol Biomarkers Prev 4:589–594.
22. Cohen L (1987): Diet and cancer. Sci Am 257:42–48.
23. Willett WC (1990): Vitamin A and lung cancer. Nutr Rev 48:201–211.
24. Block G (1992): The data support a role for antioxidants in reducing cancer risk. Nutr Rev 50:207–213.
25. Patcht ER, Kaseki H, Mohammed JR, Cornwell DG, Davis WB (1986): Deficiency of vitamin E in the alveolar fluid of cigarette smokers. Influence on alveolar macrophage cytotoxity. J Clin Invest 77:789–796.
26. Frei B, Forte TM, Ames BN, Cross CE (1991): Gas phase oxidants of cigarette smoke induce lipid peroxidation and changes in lipoprotein properties in human blood plasma. Biochem J 277:133–138.
27. Shields PG, Bowman ED, Harrington AM, Doan VT, Weston A (1993): Polycyclic aromatic hydrocarbon-DNA adducts in human lung and cancer susceptibility genes. Cancer Res 53:3486–3492.
28. Tang D, Santella RM, Blackwood A, Young T-L, Mayer J, Jaretzki A, Grantham S, Carberry D, Steinglass KM, Tsai W-Y, Perera FP (1995): A molecular epidemiological case-control study of lung cancer. Cancer Epidemiol Biomarkers Prev 4:341–346.
29. Grinberg Funes RA, Singh VN, Perera FP, Bell DA, Young TL, Dickey C, Wang LW, Santella RM (1994): Polycyclic aromatic hydrocarbon—DNA adducts in smokers and their relationship to micronutrient levels and glutathione-S-transferase M1 genotype. Carcinogenesis 15:2449–2454.
30. Freedman LS, Schatzkin A, Schiffman MH (1992): Statistical validation of intermediate markers of precancer for use as endpoints in chemoprevention trials. J Cell Biochem 16G (Suppl):27–32.
31. Benner SE, Hong WK, Lippman SM, Lee JS, Hittelman WM (1992): Intermediate biomarkers in upper aerodigestive tract and lung chemoprevention trials. J Cell Biochem 16G (Suppl):33–38.
32. Hollstein M, Sidransky D, Vogelstein B, Harris CC (1991): p53 Mutations in human cancers. Science 253: 49–53.
33. Suzuki H, Takahashi T, Kuroishi T, Suyama M, Ariyoshi Y, Takahashi T, Ueda R (1992): p53 Mutations in non-small cell lung cancer in Japan: Association between mutations and smoking. Cancer Res 52:734–736.