Molecular Epidemiology: Insights Into Cancer Susceptibility, Risk Assessment, and Prevention

Frederica P. Perera*

Recent molecular epidemiologic research provides compelling new evidence that environmental factors are major contributors to human cancer and that their risks are strongly influenced by genetic and acquired susceptibility. In particular, molecular epidemiology has demonstrated substantial variability in biologic response to carcinogens and suggests that certain groups—such as the very young, those with predisposing genetic traits or nutritional deficits, and even certain ethnic groups—are likely to have greater risk from selected exposures than other members of the population. This work implies that major gains in prevention of cancer, which will claim more than 554,000 American lives this year, will necessitate health and regulatory policies that protect these more susceptible groups and individuals from risks of man-made and naturally occurring environmental carcinogens. The specific implication from this research is that, to be effective in prevention, risk assessments developed in support of these policies by regulatory bodies, such as the Environmental Protection Agency, should reflect the available scientific data on individual variability in both exposure and susceptibility. [J Natl Cancer Inst 1996; 88:496-509]

For more than a decade, there has been wide agreement that most cancer results from man-made and natural environmental exposures (such as tobacco smoke; chemical pollutants in air, water, food, drugs; radiation; dietary constituents; radon; and infectious agents) acting in concert with both genetic and acquired characteristics (1-5) (Fig. 1). It has been estimated that without these environmental factors, cancer incidence would be dramatically reduced, by as much as 80%-90% (4,5). Cancer risk from these environmental carcinogens is strongly influenced by many factors, including genetics, age, ethnicity, sex, immune function, pre-existing disease, and level of nutrition. Genetic predisposition acting in isolation probably explains no more than 5% of all cancers in the United States (6).

The counterpoint between the environment and host susceptibility has been elaborated in recent years through molecular epidemiologic studies of biologic markers present in human cells, tissues, and body fluids (2,7-10). As shown in Fig. 2, examples of "biomarkers" are chemical-specific genetic damage in the form of carcinogen-DNA adducts or changes in key oncogenes or tumor suppressor genes that either trigger or block cancer development. Others document genetic and acquired susceptibility traits affecting processes such as carcinogen metabolism, DNA damage, and repair. Analysis of biomarkers is increasingly being incorporated into cross-sectional, retrospective, prospective, or nested case-control studies to gain improved resolution of the risk factors and mechanisms responsible for cancer.

Molecular epidemiology has the advantage of being directly relevant to human risk, unlike animal or other experimental models that require extrapolation to humans (11). In contrast to traditional epidemiology that relies on cancer incidence or mortality as the end point, molecular epidemiology has the potential to give early warning by flagging the preclinical effects of exposure and increased susceptibility, thus signaling opportunities to avert cancer through timely intervention. Moreover, biomarker data on the distribution of procarcinogenic changes and susceptibility factors in the population can improve estimation of cancer risk from a given exposure (12).

However, molecular epidemiology is also subject to many of the limitations of epidemiology, such as the vulnerability to confounding factors that can give misleading results (9,10). As with any kind of biomonitoring and screening, it carries with it the potential for social harm [reviewed in (6,13-17)]. At the present time, many biomarkers can be useful in assessing exposure, dose, and potential risk for a group or population; most are not, however, sufficiently characterized in terms of their ability to predict disease for use in screening, diagnosis, or quantitative estimation of individual risk (13,18). Even in the case of relatively well-characterized, rare, predisposing genetic traits, such as the breast and colon cancer genes, the many interactions be-

*Correspondence to: Frederica P. Perera, Dr.P.H., Columbia University School of Public Health, Division of Environmental Health Sciences, 60 Haven Ave., B-1, New York, NY 10032.
See "Notes" section following "References."
tween biologic host factors and the environment complicate individual risk assessment. Another consideration is that most of the common predisposing traits are specific to certain exposures and types of cancer. Although routine screening for cancer susceptibility is not justified, the substantial body of molecular epidemiologic and other data on susceptibility now in existence can be helpful in shaping cancer risk assessment, public health, and environmental policy.

Renewed efforts are needed to prevent cancer: Each year over half a million people die of the disease in the United States (more than 1500 people every day) and more than 1 million new cases are diagnosed (19). The pain and suffering inflicted by cancer cannot be quantified, but the immediate economic cost of cancer is $104 billion annually in the United States (19). Prevention of only 20% of the 1,359,000 new cases of cancer diagnosed every year would lead to at least 271,000 fewer cases each year in the United States alone (19). The relative contributions of “man-made” versus “natural” carcinogens are controversial and are yet to be resolved (20, 21). Thus, a practical approach to the problem of cancer is to combine active intervention on both fronts: modifying hazardous lifestyles (smoking, a diet high in fat and low in protective antioxidants and fibers) and reducing or eliminating involuntary exposure to carcinogenic pollutants that pose substantial risks.

Unfortunately, prevention is not well served by current methods to assess environmental risks. Although sometimes characterized as overly conservative, risk assessment methods currently used by government agencies, such as the Environmental Protection Agency, may underestimate the risk of certain environmental carcinogens because of the inherent assumption that all individuals in a population have the same biologic response to a specified dose of a cancer-causing agent (22). The
result can be regulatory and health policies aimed at protecting the "average American," ignoring the sizable, more vulnerable, fraction of the population. While the assumption of population homogeneity has long been in question, considerable quantitative information on the range of susceptibility within populations now shows that it is invalid. For example, extrapolating from the observed human variation in only a few indices of susceptibility (carcinogen metabolism and genetic damage), at a given exposure, one person might be tenfold to several hundredfold more susceptible to cancer than another (6,23,24). This information has obvious implications for cancer risk assessment that seem especially timely given recent congressional proposals to base major decisions about protecting health and the environment primarily on risk assessment and cost–benefit analysis (25).

This review discusses representative examples of molecular epidemiologic research during the last 10 years into the nature of susceptibility to common environmental carcinogens. These are viewed within the context of related experimental and epidemiologic data on carcinogenesis. A brief summary of recent evidence on molecular mechanisms in carcinogenesis provides a framework for an in-depth discussion of molecular and epidemiologic evidence of the role of the environment in cancer. The sources of variation in individual susceptibility to environmental carcinogens are then described in detail, with illustrative examples from the recent literature (Table 1). The final section discusses some of the implications of individual variability for current policies concerning risk assessment and cancer prevention.

**Mechanisms in Carcinogenesis**

Until fairly recently, the carcinogenic process was pictured as an orderly progression of the cell through three distinct stages: initiation by exposure to genotoxic agents; tumor promotion by agents that stimulated the initiated cell to proliferate and expand clonally to form a benign tumor; and progression, in which the accumulation of additional genetic damage in the expanding population of initiated cells caused the tumor to become malignant. This simplified model has been modified by the discovery

---

**Table 1. Known or potential biologic susceptibility factors in cancer**

<table>
<thead>
<tr>
<th>Type</th>
<th>Factor*</th>
<th>Prevalence</th>
<th>Cancer risk†</th>
<th>Mechanism</th>
<th>Reference Nos.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic factors‡</td>
<td>Li–Fraumeni</td>
<td>&lt;0.1%</td>
<td>Breast, other</td>
<td>Loss/inactivation</td>
<td>(26,27)</td>
</tr>
<tr>
<td>Inherited syndromes (rare)</td>
<td>Rb</td>
<td>&lt;0.1%</td>
<td>Retinoblastoma</td>
<td></td>
<td>(28-30)</td>
</tr>
<tr>
<td></td>
<td>Wilms' tumor</td>
<td>&lt;0.1%</td>
<td>Bladder</td>
<td></td>
<td>(31)</td>
</tr>
<tr>
<td></td>
<td>BRCA1</td>
<td>&lt;0.1%</td>
<td>Breast</td>
<td></td>
<td>(32,33)</td>
</tr>
<tr>
<td></td>
<td>FAP</td>
<td>&lt;0.1%</td>
<td>Colon</td>
<td></td>
<td>(34)</td>
</tr>
<tr>
<td></td>
<td>HNPCC</td>
<td>&lt;0.1%</td>
<td>Colon</td>
<td></td>
<td>(34)</td>
</tr>
<tr>
<td></td>
<td>XP</td>
<td>&lt;0.1%</td>
<td>Skin</td>
<td></td>
<td>(34)</td>
</tr>
<tr>
<td></td>
<td>AT homozygote</td>
<td>&lt;0.1%</td>
<td>Breast, other</td>
<td></td>
<td>(32,37)</td>
</tr>
<tr>
<td></td>
<td>AT heterozygote</td>
<td>1%</td>
<td>Breast, other</td>
<td></td>
<td>(37)</td>
</tr>
<tr>
<td></td>
<td>CYP1A1</td>
<td>10% (C)‡</td>
<td>Lung, other</td>
<td>Altered metabolism</td>
<td>(23)</td>
</tr>
<tr>
<td></td>
<td>CYP2D6</td>
<td>90%-95% (C)‡</td>
<td>Lung</td>
<td>Altered metabolism</td>
<td>(23,38)</td>
</tr>
<tr>
<td></td>
<td>GST</td>
<td>50% (C); 35% (A-A)‡</td>
<td>Lung, bladder</td>
<td>Decreased detoxification</td>
<td>(39,40)</td>
</tr>
<tr>
<td></td>
<td>NAT2 (slow)</td>
<td>50%-60% (C); 30%-40% (A-A); 14% (A-A)</td>
<td>Bladder, breast</td>
<td>Decreased detoxification</td>
<td>(41-44)</td>
</tr>
<tr>
<td></td>
<td>O²-Alkyldeoxyguanosine</td>
<td>18% (A–A); 3% (C)‡</td>
<td>Lung, other</td>
<td>Inefficient DNA repair</td>
<td>(45)</td>
</tr>
<tr>
<td></td>
<td>Hras-1 VTR</td>
<td></td>
<td>Lung, breast, other</td>
<td>Unknown</td>
<td>(46-49)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Differing prevalence of genotypes (see above)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Young or old at age at exposure</td>
<td>10% (≤5 y); 7% (15–19 y); 13% (&gt;65 y)</td>
<td>Breast, lung, other</td>
<td>Decreased detoxification, DNA repair, and immune function</td>
<td>(50-54)</td>
</tr>
<tr>
<td>Gender</td>
<td>Hormonal factors</td>
<td>51% (female)</td>
<td>Breast</td>
<td>Deregulation of growth and differentiation via receptor binding</td>
<td>(55)</td>
</tr>
<tr>
<td>Pre-existing health impairment</td>
<td>Immunologic impairment</td>
<td>Various</td>
<td>Liver, lung, breast, cervical, other</td>
<td>Decreased immune function, altered metabolism, detoxification; deregulation of growth and differentiation</td>
<td>(53,56) (5,57-59)</td>
</tr>
<tr>
<td>Pre-existing disease</td>
<td>Nutritional factors</td>
<td>Nutritional deficits</td>
<td>Lung, breast, cervical, other</td>
<td></td>
<td>(60,61)</td>
</tr>
</tbody>
</table>

*BRCA1 = breast cancer susceptibility; FAP = familial adenomatous polyposis; HNPCC = hereditary nonpolyposis colon cancer; XP = xeroderma pigmentosum; AT = ataxia telangiectasia; GST = glutathione-S-transferase; and NAT2 = N'-acetyltransferase 2.
†Risks reported in some, but not all, studies. Risks are strongly dependent on exposure.
§Effect may be modified by environmental exposures, acquired susceptibility, or gene–gene interactions.
§C = Caucasian; A-A = African-American; and As = Asian.
1% of U.S. population.

Journal of the National Cancer Institute, Vol. 88, No. 8, April 17, 1996
that cancer results from a succession of genetic and epigenetic events whose order may vary [reviewed in (5,62-66)]. Carcinogens are now understood to be remarkably versatile, able to derail gene function by inducing mutations or by disrupting gene expression or both (62). So-called "nongenotoxic" agents, such as chlorinated organic compounds, hormones, and asbestos, are known to indirectly damage the genes via a number of different mechanisms, including alterations in gene expression and oxidant formation (63,65,67-69).

The current paradigm holds that cancer results from the accumulation of changes in the structure or expression of certain key genes by mechanisms as varied as point mutation induced by carcinogen–DNA binding, gene amplification, translocation, chromosomal loss, somatic recombination, gene conversion, or DNA methylation (6,62,64,66,70,71). At the center of the paradigm are the oncogenes and tumor suppressor genes that code for proteins serving as "relays" in the regulatory circuitry of the cell (63). Damage to these target genes can result in altered protein products or abnormal amounts of normal proteins, leading to deregulation of cell growth and differentiation. The carcinogenic process involves the accumulation of genetic changes that can be facilitated by the many susceptibility factors shown in Table 1. Fig. 3 summarizes the current paradigm of carcinogens.

**Environmental Factors in Cancer**

The context of this discussion of environment–susceptibility interactions is the substantial body of epidemiologic and molecular epidemiologic research that has clearly established the central role of external environmental factors in cancer (1,5,72). The known environmental causes of cancer include chemicals in cigarette smoke, the diet, the workplace, and the general environment, as well as radiation, therapeutic agents, and viruses (3,5). Additive and even synergistic interactions between various factors occur commonly. Oxidative agents generated by natural endogenous metabolic reactions and by foreign, exogenous chemicals are also increasingly viewed as important in carcinogenesis (5).

Molecular epidemiology has taken advantage of the fact that many environmental agents leave characteristic fingerprints on DNA when they bind covalently to nucleotides in DNA to form chemical–DNA "adducts." If unrepaired prior to cell replication, the adducts cause mutations, increasing the probability of cancer development. Although circumstantial in nature, when consistent with other epidemiologic and molecular data, this evidence enables linkages between specific exposures and cancer. Chemical-specific DNA damage is considered both an indicator of the molecular dose/potential risk of carcinogens and of interindividual variation in activating and detoxifying/repair pathways that determine risk from genotoxic environmental carcinogens.

Since 1982, with our first report of the occurrence of carcinogen–DNA adducts in human cells in vivo (7), many different carcinogen–DNA adducts have been detected in humans. Carcinogen adducts with proteins such as hemoglobin have also been widely measured as a proxy for DNA damage (73). The chemicals detected on DNA or protein include such diverse agents as ethylene oxide, 4-aminobiphenyl (4-ABP), PAHs, aflatoxin B1 (AFB1), nitrosamines, and alkylating agents to metals, asbestos fibers, steroid hormones, and aflatoxin B1 (AFB1).

To illustrate, a variety of molecular epidemiologic studies have demonstrated that men, women, children, and newborn infants exposed to carcinogenic PAH in tobacco smoke, air pollution, or the diet have significantly higher levels of carcinogen–DNA or carcinogen–protein adducts in their blood cells than those without exposure (74-82). In workers and residents exposed to ambient pollution, PAH–DNA and other aromatic adducts were correlated with gene and chromosome mutation in the same individuals (80,83). Thus, the adducts provided a molecular link between environmental exposures and effects more closely associated with cancer risk. Similarly, the hemoglobin of smokers contains significantly more of the tobacco carcinogen 4-ABP than hemoglobin of nonsmokers (84,85). Overall, the adduct data show a dose–response relationship for these genotoxic carcinogens, with no apparent threshold. Even at the lowest levels of exposure, a certain fraction of the population manifests molecular damage.

Adducts have gained relevance as a potential risk marker from the recent finding that patients with lung cancer had
markedly higher PAH-DNA adduct levels than people without cancer, after taking into account differences in the amount of smoking, dietary, and other PAH exposures (86). The study suggests that persons who form high levels of adducts in response to PAH have a sevenfold greater risk of developing lung cancer than those who sustain less genetic damage. Although, in general, case-control studies have the limitation that biomarkers may be influenced by the disease, these results are consistent with the observation that about one in 10 heavy smokers develops lung cancer (87).

Other molecular epidemiologic studies have linked DNA damage from AFB1 and 4-ABP to cancer risk. In a Chinese population, the presence of AFB1-DNA adducts or AFB metabolites in urine conferred a threefold to fourfold risk of liver cancer that increased to almost 60-fold in combination with hepatitis B infection (57,88). Similarly, levels of 4-ABP-protein adducts in smokers of light versus dark tobacco were proportional to bladder cancer risk for these two groups of smokers (89).

In every case, the levels of DNA damage vary considerably between persons with similar exposure. With respect to PAH-DNA adducts, the range observed in peripheral leukocytes from many different populations using the same analytical method was 30- to 50-fold (80,82,90,91). The observed variability reflects a combination of true biologic variability, unaccounted for differences in exposure, within-person variability, and laboratory variation, with a significant fraction apparently contributed by biologic factors. Research discussed below is elucidating the biologic sources of variation in response to environmental carcinogens.

Characteristic patterns of mutations in tumors also provide valuable, albeit circumstantial, molecular evidence of the importance of environmental carcinogens. Because mutations of the p53 tumor suppressor gene and the ras oncogene are common events in human cancer, they can be useful reporters of carcinogens that have been involved in the development of cancer (41,62,92,93). In a number of cases, the patterns of p53 and ras mutations in human tumors have been consistent with the types of DNA adducts and with the mutations induced experimentally by the specific environmental carcinogens under investigation [reviewed in (62,94)]. Thus, there is a growing body of research in which mutational spectra in p53 or ras have either corroborated known environmental causes of cancer or have generated hypotheses as to others. For instance, although p53 mutations occur in 40%-50% of colon, lung, breast, stomach, esophagus, liver, brain, and lymphoid cancers, the p53 mutational spectrum differs by type of cancer. The most common p53 mutations in cancers of the lung, breast, stomach, and liver are G→T transversions, consistent with DNA adducts of bulky carcinogens (such as PAHs) or with oxidative damage (95). In fact, G→T transversions in p53 in non-small-cell carcinomas of the lung have been associated with cigarette smoking, a major source of both PAHs and oxidants (96). The predominance of the same p53 G→T mutation in breast cancer suggests the contribution of an exogenous genotoxic carcinogen(s) (97) and is consistent with the recent detection in breast tumor tissue of DNA damage related to exposure to cigarette smoke and aromatic compounds, such as PAHs (98,99). In lung adenocarcinomas, Ki-ras mutations in codon 12 are also associated with smoking and are largely the same G→T transversions (100,101). One of the PAHs, benzo[a]pyrene (BP), induces the same G→T mutation in the Ki-ras oncogene in lung tumors of mice (102). By contrast, a different p53 mutational hotspot (codon 249 AGG to ATG transversion) is seen in lung tumors from radon-exposed uranium miners, which distinguishes them from lung tumors caused by smoking alone (103).

Distinct signal patterns of p53 mutations are also seen in liver tumors from areas where AFB1 and hepatitis B virus are risk factors (104-106), in liver tumors in vinyl chloride-exposed workers (107), in colon malignancies (62), and in UV-induced skin tumors (108).

Although analysis of mutational spectra has already become an extremely valuable tool in establishing causation between environmental risk factors and cancer, the spectra are not unique to a single agent and so must be carefully interpreted in light of other human and experimental data.

Interindividual Variation in Susceptibility

Table 1 summarizes the prevalence of susceptibility factors, the type of cancer affected, and the known or putative mechanism involved.

Genetic Factors

Rare Genetic Disorders

Genetic factors participate to varying degrees in the complex interplay between the individual and his or her environment (13). These range from rare, highly penetrant, dominant mutations to more common genetic variants that influence the individual's response to environmental exposures. Examples of dominant genetic disorders that confer very high inherent susceptibility include the rare Li-Fraumeni syndrome in which persons inheriting a germline deletion of one allele of the tumor suppressor p53 gene are at extraordinarily high (almost 100%) risk of breast and other cancers (26,27). Another disorder is familial bilateral retinoblastoma, which occurs in about five of 100,000 children born in the United States (28) and results from an inherited deletion of one allele of the Rb tumor suppressor gene and a subsequent inactivation of the other allele (29,30,109). Bilateral Wilms' tumor, a rare pediatric kidney tumor, similarly involves inactivation of a tumor suppressor gene (WT1) (31). Women with an inherited mutation in the breast cancer susceptibility gene, BRCA1, are at greater than 70% risk of developing breast cancer by age 70 years (32,33). Two types of colon cancer, familial adenomatous polyposis (FAP) and hereditary nonpolyposis colon cancer (HNPPC), are due to germline mutations thought to affect the APC tumor suppressor gene and a DNA mismatch repair gene, respectively (34). The lifetime risks of developing colon cancer are 50% for offspring of APC carriers and 80% for persons with the HNPPC gene (34). As with the previous examples, although devastating to the affected individuals, inheritance of these mutated genes explains only a small proportion (perhaps 5%) of the overall incidence of the disease.

Other rare inherited disorders lead to increased susceptibility through defective DNA repair, inefficient proofreading during
DNA replication, or chromosomal instability, all of which enhance the rate of mutation. These include xeroderma pigmentosum (XP), ataxia telangiectasia (AT), Bloom’s syndrome, and Fanconi’s anemia (5,6). The AT syndrome occurs rarely (one in 40,000 persons) but confers a 100-fold risk of cancer. Although the associated risk to individuals who are heterozygous (have inherited a single defective AT allele) is far lower than that of homozygotes, this genotype occurs much more commonly (14 in 1000 persons or about 1%) and may be responsible for 7% or more of breast cancers in the United States (32,36,37).

From research on hereditary cancer has come the understanding that sporadic nonhereditary forms of cancer often involve the same genetic alterations (26). Indeed, the p53 gene is mutated in about half of the common tumors in the United States (6). Thus, by an unlucky stroke, through the inheritance of altered genes, some individuals are placed on the “fast track” to cancer, while others apparently acquire the same alterations as somatic mutations during their lifetime (26).

More Common Genetic Traits

Metabolism and detoxification. Given equal exposure to the same carcinogen, individuals will vary in their internal processing of the agent, depending on genetic background, acquired characteristics, and other past or ongoing exposures. Although their relationship to risk is less clear-cut and dramatic than that of the inherited genetic alterations just described, certain, more common genetic traits that substantially control the metabolic activation or detoxification of carcinogenic chemicals to their DNA-damaging intermediates appear to be important determinants of risk (23,110-112). The superfamily of cytochrome P450 phase I enzymes catalyzes the oxidative metabolism of most endogenous chemicals (e.g., steroids, fatty acids) and exogenous chemicals (e.g., mycotoxins ingested in the diet, drugs, and environmental contaminants, such as PAHs and aromatic amines). While this housekeeping process converts them to water-soluble, readily excretable forms, it also creates high-energy electrophilic intermediates. In fact, most genotoxic chemical carcinogens are not intrinsically reactive but require metabolic conversion to DNA-binding intermediates (62,65,113,114).

Many of the P450 genes are known to exist in variant forms or polymorphisms that have differing activities. One of these, the P450 CYP1A1, catalyzes the oxygenation of PAHs such as BP, thereby producing certain metabolites that are highly reactive with DNA. CYP1A1 is induced by cigarette smoke and environmental chemicals, such as dioxins and PAHs (23,113). Individuals vary widely in the extent to which CYP1A1 is induced by such exposures, with 10% of the Caucasian population being highly inducible (23). The level of CYP1A1 inducibility varies by 20-fold in human liver (116), while activity of the CYP1A1 enzyme differs by more than 50-fold in human lung tissues (117). A number of molecular epidemiologic studies have convincingly correlated increased risk of lung cancer in smokers with high CYP1A1 inducibility (118,119).

In Japanese, a threefold increase in lung cancer risk was associated with a polymorphism in the CYP1A1 gene, known as Msp I (120), with the greatest lung cancer risk in light smokers (121). The Msp I polymorphism is itself linked to an exon 7 mutation in the CYP1A1 gene, resulting in an altered protein having the amino acid valine in place of isoleucine (Ile-Val). Experimentally, the valine-type protein is almost twice as active in metabolizing BP as the isoleucine form, pointing to a possible mechanism by which this genetic variant affects risk (121). Among healthy U.S. volunteers, the variant CYP1A1 exon 7 genotype was associated with a threefold elevation in enzyme activity (122). Studies of the relationship between the CYP1A1 exon 7 polymorphism and DNA damage have yielded conflicting results. However, cigarette smokers of one to two packs per day with the exon 7 (Ile-Val) genotype had significantly (more than twofold) higher levels of PAH-DNA adducts in white blood cells than smokers without the CYP1A1 variant, suggesting that the trait may increase cancer risk by facilitating PAH activation and binding to DNA (91).

Genetic variation in other cytochrome P450s can also be important in modulating cancer risk. CYP1A2 metabolizes amines, heterocyclic amines, and aflatoxins. A polymorphism in the CYP1A2 gene, in combination with a variant of the N-acetyltransferase 2 (NAT2) gene, has been linked to risk of colon cancer, especially in conjunction with a diet rich in food mutagens (123). In human liver, there is a greater than 40-fold variation in expression of the CYP1A2 gene (116). Another of the P450s, CYP2E1, metabolizes N-nitrosamines, butadine, benzene, and carbon tetrachloride (124). Although its relationship to cancer risk has not been demonstrated (125), the activity of CYP2E1 in humans varies by 50-fold (124). The P450 CYP2D6 also differs among humans: 90%-95% of the U.S. Caucasian population are efficient metabolizers with 10- to 200-fold higher rates of metabolism than poor metabolizers (23). Although not all studies agree, some have indicated that smokers who are efficient CYP2D6 metabolizers are at fourfold to sixfold greater risk of lung cancer (38). The substrate for CYP2D6 is not known but may be the tobacco-specific nitrosamine 4-(methyl-nitrosamino)-1-(3-pyridyl)-1-butanone (NNK). In lung biopsies, DNA adducts related to N-nitrosamines were found to be associated with CYP2D6 genotype and to a lesser degree with CYP2E1, mainly in individuals with low exposure to tobacco smoke (126).

In contrast to the phase I P450-activating enzymes, phase II enzymes (epoxide hydrolase, glutathione S-transferase, N-acetyltransferase, and sulfortransferase) generally detoxify carcinogenic metabolites by conjugating them with glucuronide, glutathione, or sulfate to produce hydrophilic products that are readily excreted. The balance between phase I and II enzymes determines the molecular dose of carcinogens, thereby substantially influencing cancer risk. For example, glutathione S-transferase M1 (GSTM1) detoxifies a number of reactive, electrophilic substances, including the carcinogens PAHs, ethylene oxide, and styrene (39). Approximately 50% of Caucasians have a deletion in the GSTM1 gene (40). The null GSTM1 genotype or its phenotypic expression has been quite consistently (although not uniformly) associated with increased risk of bladder and lung cancers (39-41,127,128). One study (40) suggested that as much as 25% of all bladder cancers may be attributable to the GST null genotype in conjunction with smoking. As an illustration of the importance of gene–environment interactions, persons with the null genotype had little risk
of bladder cancer in the absence of environmental exposure. Risk increased by almost twofold with any exposure to tobacco smoke and by sixfold in heavy smokers (>50 pack-years) compared with unexposed persons without the deletion (40). In lung biopsies, PAH-DNA adducts were inversely related to the GSTM1 null genotype (126), although a case-control study of lung cancer found that PAH-DNA adducts in white blood cells and GSTM1 deletion were independent risk factors for lung cancer (129).

A second phase II enzyme, \( \text{N-acetyltransferase 2 (NAT2)} \), is a noninducible liver enzyme that deactivates carcinogenic aromatic amines via \( \text{N-acetylation} \). Fifty percent to 60% of Caucasians and 30%-40% of African-Americans are slow acetylators (42,43). Some, but not all, studies indicate that NAT2 slow acetylators are at higher risk of bladder cancer due to environmental exposures, such as 2-naphthylamine and 4-ABP (41). At low levels of exposure to tobacco smoke, slow acetylators had approximately twofold higher levels of 4-ABP adducts on hemoglobin than rapid acetylators, whereas the effect of NAT2 on 4-ABP adducts was not seen at higher levels of exposure (85). A recent study (44) indicates that women who smoke and have a slow-acting version of NAT2 may be eight times more likely to develop breast cancer than women who do not smoke. While NAT2 is involved in detoxification of aromatic amine bladder carcinogens, it apparently plays an opposite role in colon cancer through the metabolic activation of the carcinogenic heterocyclic amines. Thus, ironically, persons who are NAT2 rapid acetylators appear to be at increased risk of colon cancer (123).

In human cancer, not only do multiple genes modulate the effects of the environment, but interactions between genes can result in a greater than additive effect on risk. For example, in a Japanese study, individuals with the combination of CYP1A1 (Val/Val) and GSTM1 null genotypes were at a remarkably high risk of developing lung cancer (approximately fivefold overall; eightfold for squamous cell carcinoma) compared with individuals without either trait (121). Other studies of GST (130) or CYP1A1 alone (131) or in combination (132) in Caucasian populations have not shown this strong relationship with cancer risk, possibly because of ethnic differences in prevalence or the extent to which \( Msp I \) is linked to a functional mutation.

Some research suggests greater cancer risk from certain genetic traits at lower levels of exposure to carcinogens (CYP1A1 \( Msp I \) and lung cancer in Japanese; NAT2 and bladder cancer) (85,121). While this was not seen with GSTM1 and bladder cancer (40), it is plausible that at higher exposure the effects of certain genetic traits are overpowered by the environmental influence (85).

**DNA repair.** Decreased efficiency or fidelity in repair, in the absence of high-risk syndromes such as AT and XP, has been linked to increased cancer risk (5,112,133). In the general population, the molecular machinery to repair DNA damage and thereby block mutation is variable from person to person. Genetic factors, environmental exposures (such as smoking and certain chemicals), and physiologic factors can all decrease an individual’s ability to repair DNA. For example, one of the DNA repair enzymes, \( \text{O}^6\text{-alkyldeoxyguanine-DNA alkyltransferase} \), reverses DNA damage formed by \( \text{N-nitroso compounds} \) commonly found in the diet, in cigarette smoke, and in other mixtures of pollutants (134,135). Humans display marked interindividual variability in the activities of this repair enzyme (as much as 180-fold) in contrast to 20% in the inbred rat (133,135). Activity of repair enzymes also varies between tissues (135). Patients with lung cancer have reduced alkyltransferase activity compared with controls, possibly reflecting inherent susceptibility (45). Activity of another repair enzyme, uracil DNA glycosylase, also varies in humans (up to 300-fold) (133). Similarly, substantial variability is evident in excision of DNA damage following exposure to UV radiation and bulky carcinogens such as \( \text{N-acetoxy-2-acetylaminofluorene (2-AAF)} \) and PAHs (65,136,137). White blood cells from first-degree relatives of individuals with cancer showed a significant reduction in DNA repair following treatment with 2-AAF compared with individuals without a family history of cancer (138). Patients with breast cancer also had deficient UV-induced DNA repair (139).

**Receptor binding.** Genetic variation in the efficiency with which external signaling molecules (ligands) are bound to receptors is a less well-studied contributor to susceptibility than are other polymorphisms that affect metabolism or repair. However, there is evidence that affinity differences in the dioxin-binding aromatic hydrocarbon (Ah) receptor, estrogen receptor, androgen receptor, and peroxisome proliferator-activated receptor may well be important determinants of risk for certain cancers (140). A polymorphism in the estrogen receptor gene in combination with a history of spontaneous abortion has been associated with breast cancer susceptibility (141).

**Conclusion**

As discussed, by themselves, high-risk, rare inherited genetic disorders are thought to explain only a small percentage of all cancers (6). More significant in terms of cancer incidence is the interaction of environmental exposures (both naturally occurring and man-made) and a variety of more common genetic variants that convey a low risk to the individual but account for a large proportion of cancer incidence in the population (6,28,41,142). Thus, it appears that the greatest opportunity to use knowledge of inherited predisposition to prevent cancer may come from recognition of genetic polymorphisms that affect individual susceptibility to exogenous or endogenous carcinogens (142).

Clearly, the role of individual inherited genetic traits is complex. With few exceptions, their effect may be modulated by gene–gene interactions, environmental exposures, health and nutritional status, and other host factors. Moreover, the same genetic trait may protect against one type of cancer but may increase the risk of others (123). Thus, in most cases it will not be possible to estimate risk based on single genetic traits alone.

**Ethnicity or Race**

The comparatively high cancer incidence and mortality rates in African-Americans compared with those in whites (19) indicate differences in exposure and/or susceptibility to carcinogens. For example, the incidence of squamous cell esophageal cancer is more than threefold higher among black men than white men (19), while incidence of lung cancer is about 50% higher in black males (143). Black women under the age of 40 years experience a higher incidence of breast cancer than white women.
under 40 years, although the situation is reversed over the age of 40 years (144). Hispanics have generally lower cancer rates than whites or African-Americans, but important differences exist between black and white Hispanics (145). Caucasians comprise about 75% of the U.S. population; African-Americans, 12%; Hispanics, 9.5%; Asians, 3%; and American Indians, 0.5% (146).

Designation by race is theoretically based on shared physical or inherited genetic traits, in contrast to ethnicity that refers to a common cultural heritage. In practice, however, it is difficult to disentangle the effect of genetic traits associated with race from socioeconomic and behavioral factors that can also affect environmental exposure or susceptibility. Nonetheless, a review of more than 60 studies provides compelling evidence that people of color and lower incomes in the United States are disproportionately exposed to certain toxic environmental pollutants, including air particulates, lead, and hazardous waste (147). Racial disparities were reported more frequently than income disparities (147).

Molecular epidemiology suggests different patterns of exposure and/or internal handling of exogenous or endogenous carcinogens across racial or ethnic groups. For example, the spectrum of p53 mutations in breast cancer appears to vary between black and white women and between Japanese and Western women, implying exposure differences (95,148). In a California study using another type of biomarker, serum levels of organochlorines (a DDT metabolite and polychlorinated biphenyls) were markedly higher in black and Asian women than in white women (149). In black smokers, levels of cotinine (a metabolite of nicotine) exceeded those in white smokers after controlling for self-reported amount of smoking (150). While both biological variability and unmeasured differences in exposure (inhalation, cigarette brands, etc.) may be involved, this finding is consistent with the higher rates of various smoking-related cancers in blacks (150). Similarly, compared with white smokers, black smokers have a higher internal dose of the tobacco-specific lung carcinogen NNK as indicated by a lower ratio of the detoxified to active metabolite in the urine, suggesting a possible explanation for their higher lung cancer risk from smoking (151).

Among the biologic factors that might explain differential risks of cancer in certain groups are variations in prevalence of inherited genetic traits, such as those affecting carcinogen metabolism and DNA repair. Ethnic or racial differences in the frequency of genetic polymorphisms probably reflect adaptive mutations, much like sickle cell trait, which confer resistance to malaria along with its detrimental effects (152). Although African-American males do not appear to smoke more than European-Americans, their incidence of lung cancer is 50% higher (143). No single gene explains this differential cancer risk, but a novel CYP1A1 Msp I polymorphism (known as the AA RFLP) is found only in African-Americans and Africans (at a frequency of 17% in African-Americans) (143). Reports have been conflicting regarding the relationship of this RFLP to adenocarcinoma of the lung (143,153). DNA damage from PAHs appears to be greater in African-Americans than in Mexican-Americans, consistent with enhanced carcinogen activation and lower DNA repair in this group (154,155). On the other hand, the so-called “at risk” GSTM1 null genotype, which has been associated with lung and bladder cancers, occurs less frequently among blacks (35%) than whites (49%) (40). Another susceptibility marker, the NAT2 slow acetylation phenotype, occurs in 40%-60% of whites, presumably placing them at greater risk of bladder cancer from certain arylamines, compared with 35% of blacks and 14% of Asians (42,43). A corresponding trend was seen in 3- and 4-ABP-hemoglobin adducts across the three groups (43). Subjects with both the GSTM1 null and slow acetylator profile had the highest adducts (156).

Although the mechanism by which a rare polymorphism, the Hras-1 variable tandem repeat (VTR), confers susceptibility is still not known, it is associated with cancers of the lung, breast, colorectum, and bladder and leukemia [reviewed in (46,47)]. The polymorphism is found more frequently in blacks (18%) than in whites (3%) (157). Consistent with other evidence that breast cancer has a different, more aggressive biology in black women (149,158), the rare H-ras VTR allele is three to six times more strongly associated with breast cancer in blacks than whites (157). Similarly, lung cancer in blacks (but not whites) is associated with the rare H-ras allele (48,49). The higher rates of bladder cancer in blacks than in whites may also be partially explained by differences in H-ras VTR.

These results indicate that no racial or ethnic group is uniformly at higher risk of cancer than another. Rather, each possesses a distinct mix of genetic traits with the potential to increase susceptibility to specific carcinogens. As summarized in Table 1, the CYP1A1 Msp I and H-ras (VTR) polymorphisms are more prevalent in blacks, while the GSTM1 null and NAT2 genotypes are more frequent in whites. Asians have a higher prevalence of CYP1A1 Msp I and a lower frequency of the NAT2 (slow) variant than Caucasians. Thus, each group has a high frequency of one or more known or suspected “at risk” genotypes relative to the others, the predicted consequences of which will vary with the specific exposure and type of cancer being considered.

Age or Stage of Development

There is mounting evidence of significant age-related differences in susceptibility to environmental toxicants. Experimental and human data indicate that, because of differential exposure and/or physiologic immaturity, infants and children may have greater risk than adults from a variety of environmental toxicants, including PAHs, nitrosamines, pesticides, tobacco smoke, and air pollution (50,159-162). On the exposure side, relative to body weight, children take in appreciably more food, water, and air (and any carcinogens contained in them) than do adults. The very young may also have uniquely high exposures from nursing or from crawling, mouthing objects, and thumb-sucking. For example, relative to body weight, nursing infants have an estimated 10- to 20-fold higher intake of dioxin than adults (162). Young children absorb and retain more of certain contaminants than adults, while at the same time they are less efficient at detoxifying them or repairing their damage.

Experiments with a number of carcinogens (including PAHs, nitrosamines, and AFB1) show that the risk of cancer is heightened if exposure begins in utero or in infancy rather than in adulthood (50,161). Molecular epidemiologic studies support the mechanism of increased internal or biologic dose in the
young. In newborns at delivery, levels of PAH–DNA adducts in the blood exceeded those in the mothers, despite the fact that the exposure to the fetus is estimated to be about 10-fold lower (50). Similarly, in young children (<2 years old), urinary levels of 1-hydroxypyrene glucuronide, an indicator of exposure to PAHs, were higher than in their mothers (163).

Human data on the effects of the atomic bomb also demonstrate that sensitivity to radiation-induced cancer is enhanced prenatally and in the early years (51). This may be due to depressed detoxification and repair, the higher rate of cell proliferation during early stages of development, and the fact that cancers initiated in the womb and in the early years have the opportunity to develop over a period of many years (50,52).

Age-related susceptibility is likely to be compounded in a significant percentage of children in the United States by the influence of poverty. Poverty in the United States is associated both with greater exposure to environmental pollutants (147) and with poor nutritional status (164). Ten percent of Americans (23 million) are children under the age of 6 years (146). Six million U.S. children younger than 6 years of age (26% of this age group) are now living in poverty: millions of them have substandard intakes of major nutrients (164,165).

Adolescence and young adulthood are also viewed as sensitive lifestages because of augmented proliferative activity in epithelial cells of certain tissues during this period of development. In survivors of the atomic bombings, women exposed during their teens had the greatest risk of breast cancer (166). Early age at starting smoking confers a higher risk of lung, bladder, and possibly breast cancers (41,167). Similarly, long-term use of oral contraceptives by young women and exposure to viral infection at an early age have been associated with enhanced risk of breast and cervical cancers (168-170). In the United States, almost 7% of the population is between 15 and 19 years of age (146).

Heightened susceptibility in the elderly has not been thoroughly explored, but is likely to reflect the progression of cancers initiated early in life as well as an age-related decrease in immune function and DNA repair efficiency (53,54). Adults older than 65 years of age constitute about 13% of the population (146).

Gender

In contrast to gender-related hormonal factors that play an obvious role in breast cancer, other sources of variation in risk due to gender have been little studied. However, research suggests that dose for dose, women may be inherently more susceptible to certain carcinogens than men. A case-control study in Canada suggests that the association between smoking and lung cancer is appreciably stronger in women than in men (171). With a history of 40 pack years of smoking, relative to lifelong nonsmoking, women smokers had a higher elevated risk (28-fold) than men (about 10-fold) (171). Women represent 51% of the U.S. population (146).

Pre-existing Health Impairment

Immunologic impairment, pre-existing disease, and nutritional deficits can increase susceptibility to carcinogens. The immunologic defenses that recognize and destroy foreign agents, tumor cells, or virally infected cells vary with age, being generally less efficient in children and in the elderly (53,56). Moreover, UV radiation, diverse environmental chemicals, and stress are also capable of suppressing the human immune system (56,172).

Various infectious diseases (viral, bacterial, and parasitic) predispose to cancer, sometimes interacting with environmental and other host factors. Illustrations of this phenomenon include hepatitis B and AFB1 in liver cancer (57) and human papillomavirus and cigarette smoking in cervical cancer (5). Noninfectious diseases, such as cirrhosis, pulmonary fibrosis, and benign proliferative disease (2.58), also increase cancer risk. Previous lung disease (asthma and chronic bronchitis) is a significant risk factor for lung cancer in nonsmokers (59), while pernicious anemia predisposes to cancer of the stomach (173). Benign proliferative breast disease raises the risk of developing breast cancer and appears to have a genetic basis (58).

Nutritional Factors

Nutritional factors, including high dietary intake of animal fat and low intake of fiber and fruits and vegetables, are risk factors for various human cancers. Experimental and epidemiologic studies have identified a multiplicity of known or potential protective agents in the diet, including antioxidants/micronutrients (vitamins A, C, and E; various carotenoids; and selenium), ellagic acid and flavanoids, benzyl isothiocyanate, and organosulfur compounds [reviewed in (5,60,61,174-176)]. For example, a large number of epidemiologic studies have shown a protective effect of fruits and vegetables rich in antioxidants/micronutrients for diverse cancers, including lung, esophageal, oral, laryngeal, cervical, and breast cancers (60,61). The micronutrients are capable of acting via a number of mechanisms to block DNA damage, mutation, and carcinogenesis by oxygen radicals, PAHs, and other chemical carcinogens [reviewed in (60,177)].

Several recent molecular epidemiologic studies indicate that heavy smokers with low plasma levels of antioxidant micronutrients are less protected against DNA damage than smokers having higher plasma concentrations of these agents. Illustrating that multiple susceptibility factors are involved in individual responses to environmental challenges, the effect of low plasma vitamin E was seen only in those smokers who had the GSTM1 null genotype (178,179). Sensitivity to mutagens was also increased in healthy individuals with low plasma levels of antioxidant micronutrients (α-carotene, total carotenoids, and ascorbic acid) (180). There is growing interest in applying biomarkers in chemoprevention to block cancer development in sensitive or “at risk” groups (177,181-184).

Cancer Prevention and Risk Assessment

Taken as a whole, the research in the biology, genetics, epidemiology, and molecular epidemiology of cancer reviewed here has a number of implications for cancer prevention: 1) Because most cancer has an environmental component, an opportunity exists to prevent the majority of the 1,359,000 new cases diagnosed every year in the United States. 2) Certain groups (those with genetic, ethnic, or gender-related traits conferring
susceptibility; the young; the elderly; and persons with pre-existing disease or immunologic or nutritional deficits) are likely to have a greater risk than other members of the population who are similarly exposed to certain carcinogens. 3) Major gains in cancer prevention will necessitate regulation, public education, and other interventions that adequately protect these groups from environmental carcinogens. 4) To be effective in cancer prevention, risk assessment methods must account to the extent possible for individual variability in exposure and susceptibility to environmental carcinogens (12,13,22,125,185,186).

Research reviewed here contradicts the silent assumption underlying most current methods of risk assessment that all individuals are uniformly vulnerable to environmental carcinogens. As illustrated in Table 1, factors known or believed to affect susceptibility are not rare events in the human population. Thus, there are few cases where the risk is likely to be confined to a few “outliers” (22,187).

Biologically based interindividual variation in only a few susceptibility factors could lead to a significant increase in population risk over that expected based on an assumption of uniform susceptibility, possibly by an order of magnitude or more (24,188). Hattis et al. (188) have calculated that the variability in metabolic activity, detoxification, and DNA repair among 95% of the U.S. population could be as high as 85- to 500-fold. Therefore, despite the common perception that risk assessment is overly conservative, the erroneous assumption of biologic homogeneity can result in a serious underestimate of risk to the overall population and especially to the most susceptible subsets or individuals. This, in turn, can lead to environmental standards and public health policies that are not adequately protective and even create or perpetuate a situation in which disproportionately high risks are borne by the more vulnerable subset of the population (22). This concern is compounded by the evidence that certain age or ethnic groups are apt to be both biologically more susceptible and disproportionately exposed to certain environmental carcinogens (189).

It is, therefore, critical that policymakers consider interindividual differences when assessing risks and setting standards for environmental contaminants (12,161) as well as in formulating educational and other programs to protect public health. Wherever possible, risk assessments should explicitly account for the range of human susceptibility among the population (12). By presenting the distribution of risks across the population, as well as the estimated risks to specific groups known or likely to be most sensitive to the exposure in question (e.g., children, women, or specific ethnic groups), risk assessment can be far more effective in shaping public policy that is both preventive and fair. As reviewed elsewhere (189), in addition to the scientific arguments underscored by molecular epidemiology, there are compelling ethical reasons for preventive policies to protect the most vulnerable members of our society.

References


(38) M. Block G, Cowden C, Rothman N. Lung cancer and CYP2D6 (the debrisoquine polymorphism): sources of heterogeneity in the proposed association. Pharmacogenetics 1995;5:S129-34.


(40) Bartsch H, Taylor LA, Petri MP. Perspectives on the risk assessment for nongenotoxic carcinogens and tumor promoters [see comment citation in Medline]. Environ Health Perspect 1991;94:231-5.


Moon RC. Comparative aspects of carotenoids and retinoids as chemopreventive agents for cancer. J Nutr 1993;123:127-34.


Notes

Supported by Public Health Service grants CA39174, CA35809, CA51196, CA66224, and CA51196 (National Cancer Institute) and ES05294 and ES06722 (National Institute of Environmental Health Sciences), National Institutes of Health, Department of Health and Human Services; by grant DAMD17-94-J-4251 from the Department of Defense; by grant 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, the Irving A. Hansen Memorial Foundation, the Gladys and Roland Hartman Foundation, the Bauman Foundation, and the March of Dimes Birth Defects Foundation.

I gratefully acknowledge the invaluable contribution of my colleagues at Columbia University and elsewhere to the research described here, and I thank Gabriela Simon-Cereijido for her expert assistance in the preparation of the manuscript. I also thank Andrew Rundle (Columbia University), the illustrator of Figs. 1 and 2.

Manuscript received October 6, 1995; revised December 26, 1995; accepted January 26, 1996.