***Eleventh Annual AACR American Cancer Society Award for Research in Epidemiology or Prevention Lecture***

# Genetic Susceptibility to Lung Cancer: The Role of DNA Damage and Repair[1,2]

**Margaret R. Spitz,[3] Qingyi Wei, Qiong Dong, Christopher I. Amos, and Xifeng Wu**

Department of Epidemiology, The University of Texas, M. D. Anderson Cancer Center, Houston, Texas 77030

## Introduction

It has long been recognized that people differ in their susceptibility to disease. Some of these differences are ascribed to the concept that heritable traits modify the effects of environmental exposures. Lung cancer epitomizes this concept. More than 80% of lung cancers are attributed to tobacco (1). However, only a fraction of smokers (~15%) will develop lung cancer in their lifetime. One of the several cellular processes that could explain this interindividual variation in risk is DRC,[4] the focus of this paper.

Genetic instability, which drives tumorigenesis, is itself fuelled by DNA damage and by errors made by the DNA repair machinery. Cancer syndromes such as XP and Bloom's syndrome support the involvement of DNA repair processes in governing the amount of genomic instability that a cell experiences (2). DNA repair is a ubiquitous defense mechanism that is critical to maintaining the integrity of the genome and repairing the damage from exposure to exogenous environmental xenobiotics, as well as to endogenous damage (*e.g.,* from oxidative metabolism) or spontaneous disintegration of chemical bonds in DNA. There is quite substantial interindividual variation in DRC within a population. At the extreme end of this spectrum are patients with XP, who have a defect in NER, and who exhibit thousand-fold increased risks of skin cancer. There is a larger subgroup with reduced DRC who are likely to be at increased cancer risk, but are phenotypically normal. The challenge for molecular epidemiological research is to be able to identify this at-risk subgroup, who could be targeted for the most intensive behavior modification changes and screening interventions.

In an extensive review of the published literature on DNA repair and susceptibility to cancer in humans, Berwick and Vineis (3) concluded that "the vast majority of studies show a difference between cancer case subjects and control subjects." They also sounded a cautionary note regarding the issues of confounding, and the need to develop molecular assays that define both the genetic defect and the repair pathways involved.

Our approach to risk assessment is multitiered, beginning with a detailed epidemiological assessment in case-control studies, followed by the application of an array of phenotypic and genotypic markers of genetic susceptibility. This article is not a review of the literature but a summary of selected published and unpublished data from our own laboratories assessing the role of DRC in lung cancer risk.

**Phenotypic Assays of DRC.** PBLs are the tissue of first choice for molecular epidemiological studies that assay DRC (4). The types of assays include: (*a*) those using a chemical or physical mutagen challenge (such as the mutagen sensitivity, Comet, and induced adduct assays); (*b*) unscheduled DNA synthesis; and (*c*) measuring cellular ability to remove adducts from plasmids transfected into lymphocyte cultures *in vitro* by expression of damaged reporter genes (the host-cell reactivation assay). This last assay is a direct measure of repair kinetics (4), unlike the cytogenetic assays that only indirectly infer DRC from cellular damage remaining after mutagenic exposure and recovery (3), and as such probably reflecting general and nonspecific impairment of the DNA repair machinery.

**Host Cell Reactivation Assay.** Although there are many assays that measure the efficiency of the multiple steps of excision repair individually, the ability to test the whole pathway is often needed for population studies, in which time, cost, and repeatability of measurements are major concerns. Therefore, measuring the expression level of damaged reporter genes using host-cell reactivation is the assay of choice. This assay uses undamaged cells, is relatively fast, and is an objective way of measuring intrinsic cellular repair. In the assay, lymphocytes are transfected with damaged nonreplicating recombinant plasmid harboring a *CAT* gene (*pCMVcat*; Ref. 4). To study tobacco-related cancers, the mutagen challenge is benzo(*a*)pyrene, a major constituent of tobacco smoke (5). Activated benzo(*a*)pyrene, BPDE, can irreversibly damage DNA by covalent binding or oxidation (6); such BPDE-DNA adducts are repaired by the NER pathway that eliminates the widest variety of damage to the human genome, including UV-induced photoproducts, bulky monoadducts, cross-links, and oxidative damage (7). Experimental conditions produce at least one BPDE-DNA adduct per plasmid, such that transcription of the *CAT* reporter gene is completely blocked, but without inducing conformational changes in the DNA. This is important, because conformational change of the plasmid could reduce the transfection rate. Because even a single unrepaired DNA adduct can effectively block CAT transcription (8), any measurable CAT

Received 2/27/03; revised 4/30/03; accepted 5/13/03.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] This study was supported by National Cancer Institute Grants CA 55769 and CA 86390 (to M. R. S.).
[2] Presented at the 93rd Annual Meeting of the American Association for Cancer Research in San Francisco, CA 2002. Dedicated to the memory of T. C. Hsu, Ph.D. (1917–2003), our revered teacher, friend, and collaborator.
[3] To whom requests for reprints should be addressed, at Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. E-mail: mspitz@mdanderson.org.
[4] The abbreviations used are: DRC, DNA repair capacity; XP, xeroderma pigmentosum; BPDE, benzo(a)pyrene-diol-epoxide; CAT, chloramphenicol acetyltransferase reporter gene; LUC, luciferase; b/c, break per cell; OR, odds ratio; CI, confidence interval; BER, base excision repair; NER, nucleotide excision repair; PBL, peripheral blood lymphocyte; SCE, sister chromatid exchange.

*Table 1* Risk estimates for lung cancer associated with DRC

| DRC (%) | Cases | Controls | Adjusted OR[a] (95% CI) |
|---|---|---|---|
| Overall | | | |
| >8.06 | 314 | 340 | Ref. |
| ≤8.06 | 450 | 337 | 1.50 (1.2–1.9) |
| Quartiles | | | |
| >10.46 | 126 | 170 | Ref. |
| 8.06–10.46 | 188 | 170 | 1.51 (1.1–2.1) |
| 6.25–8.06 | 221 | 167 | 1.85 (1.4–2.5) |
| ≤6.25 | 229 | 170 | 1.94 (1.4–2.6) |
| | | | $P < 0.001$(trend) |

[a] Adjusted by age, gender, and smoking status.

activity will reflect the ability of the transfected cells to remove BPDE-induced adducts from the plasmids.

Details of the assay have been reported previously (4, 9). Briefly, the frozen lymphocytes of each patient are thawed and processed to ensure a cellular viability of >80%. The cells are then stimulated so that they take up the plasmids (10, 11). The number of viable, large lymphoblasts in the culture for each sample is counted to calculate the blastogenic rate [(number of lymphoblasts/number of lymphocytes stimulated) × 100]. Duplicate transfections with either untreated plasmids or BPDE-treated plasmids are always performed. CAT activity is assayed by adding chloramphenicol and [$^3$H]acetyl-CoA, and measuring the production of [$^3$H]monoacetylated and [$^3$H]diacetylated chloramphenicols with a scintillation counter. DRC is reported as the ratio of the radioactivity of cells transfected with BPDE-treated plasmids to that of cells transfected with untreated plasmids. Assuming that the transfection efficiencies of BPDE-treated and untreated plasmids are equal (12), this ratio reflects the percentage of damaged CAT reporter genes repaired in test lymphocytes transfected with BPDE-treated plasmids.

Both lymphocytes and skin fibroblasts from patients who have basal cell carcinoma but not XP have lower excision-repair rates than individuals without cancer (13). Consequently, the repair capacity of lymphocytes can be considered a reflection of the overall repair capacity of an individual and a quantitative measurement of the DRC of the host cells.

We have applied this assay in lung cancer. In both an initial pilot study of 51 patients and 56 frequency-matched controls (14), and in a subsequent large hospital-based case-control study of 316 newly diagnosed lung cancer patients and 316 controls (9), statistically significantly lower DRC was observed in case patients compared with control subjects. This suboptimal DRC was associated with a >2-fold increased risk of lung cancer. Compared with the highest DRC quartile in the control subjects, DRC in the second, third, and fourth quartiles was associated with adjusted risks for lung cancer of 1.8, 2.0, and 4.3, respectively ($P_{trend} < 0.001$). Case patients, who were younger at diagnosis (<60 years), were female, were lighter smokers, or who reported a family history of cancer in a first-degree relative, exhibited the lowest DRC and the highest lung cancer risk among their respective categories, suggesting that these subgroups may be especially susceptible to lung carcinogenesis.

We have recently expanded this study and confirmed these findings in a much larger case-control analysis with a more-than-doubled sample size. The mean DRC ± SD for 764 lung cancer cases was 7.82% (±2.82) compared with 8.79% (±3.87) for the 677 controls ($P < 0.001$). The adjusted univariate risk estimate associated with suboptimal DRC (defined as below the control median) was 1.50 (1.2–1.9; Table 1). When evaluated

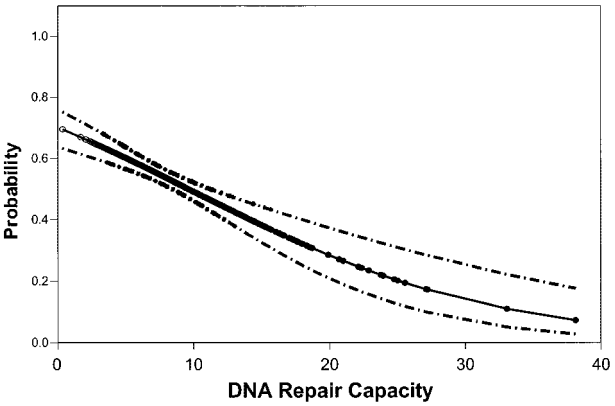


*Fig. 1.* Probability of lung cancer: univariate model for DRC with 95% CIs.

by quartiles of DRC, the risks increased with diminishing repair capacity. The respective ORs and 95% CIs for the second, third, and fourth quartiles relative to the most efficient repair stratum were 1.51 (1.1–2.1), 1.85 (1.4–2.5), and 1.94 (1.4–2.6), respectively ($P_{trend} < 0.001$; Table 1). Whereas these risk estimates are lower than those reported above for the smaller studies, they are all statistically significant, with tighter CIs, and the trend is similar. Fig. 1 depicts the relationship between the linear predictor (DRC) with its associated 95% confidence limits and the probability of lung cancer from a univariate linear logistic model. As repair capacity diminishes, the probability of being a case increases.

DRC was significantly lower in women (7.46%) compared with men (8.15%) among the cases ($P < 0.001$), and the trend for controls was similar ($P = 0.047$; Table 2). This observation, that women have significantly lower DRC than men, is consistent with our previous findings in studies of lung cancer (14) and basal cell carcinoma, in which women exhibited lower DRC than men and tended to have a much higher risk of nonmelanoma skin cancer if they also reported exposure to sunlight (13). This finding is also consistent with some reports (15, 16) that women have a higher lung cancer risk than men.

These data also show consistently that younger case patients exhibit lower DRC than their matched control subjects, suggesting that other undetected confounders or genetic predisposing factors may also be important in this subgroup. As we have reported previously, DRC was highest among currently smoking cases (7.96%) and controls (9.08%) compared with former (7.85% and 8.68%) and never smoking cases and controls (7.26% and 8.43%, respectively; Table 2). Similarly, we demonstrated a trend for more efficient DRC in heavier *versus* lighter smokers both among the cases (8.02% *versus* 7.73%) and the controls (9.20% *versus* 8.52%), although the difference was statistically significant only among the controls ($P = 0.04$).

That heavy smokers among both case patients and control subjects tended to have more proficient DRC than lighter smokers, suggests that cigarette smoking may, in fact, stimulate DRC in response to the DNA damage caused by tobacco carcinogens. Such an adaptation would be consistent with a baseline DRC that can be mobilized on increased demand for repair (17, 18). An inducible enhancement of DRC may explain why no age-related decline in DRC was evident in the control subjects, most of whom were current or former smokers. The finding of an increase in DRC with smoking is also consistent with our previous finding (19) that, in head and neck cancer patients,

*Table 2* DRC (%) for cases and controls by selected subgroups

| Subgroup | Cases | | | | Controls | | | |
|---|---|---|---|---|---|---|---|---|
| | *n* | Mean | SD | *P* | *n* | Mean | SD | *P* |
| Age (years) | | | | | | | | |
| ≤63 | 409 | 7.69 | 2.70 | Ref | 360 | 8.55 | 3.85 | Ref. |
| >63 | 355 | 8.00 | 2.94 | 0.096 | 317 | 9.05 | 3.88 | 0.091 |
| Gender | | | | | | | | |
| Female | 362 | 7.46 | 2.78 | Ref. | 321 | 8.48 | 3.50 | Ref. |
| Male | 402 | 8.15 | 2.82 | <0.001 | 356 | 9.06 | 4.16 | 0.047 |
| Smoking Status | | | | | | | | |
| Never | 97 | 7.26 | 2.45 | Ref. | 89 | 8.43 | 3.29 | Ref. |
| Former | 351 | 7.85 | 2.75 | 0.054 | 352 | 8.68 | 3.74 | 0.563 |
| Current | 316 | 7.96 | 2.98 | 0.021 | 236 | 9.08 | 4.23 | 0.142 |
| Pack-years | | | | | | | | |
| ≤42 | 268 | 7.73 | 2.72 | Ref. | 308 | 8.52 | 3.59 | Ref. |
| >42 | 399 | 8.02 | 2.95 | 0.204 | 280 | 9.20 | 4.28 | 0.040 |

DNA repair gene expression was increased in the heavy smokers among both case patients and control subjects relative to the lighter smokers. Therefore, inefficient repair response or inability to mobilize DRC on tobacco exposure may lead to accumulation of genetic damage. Indeed, case patients who were lighter smokers appeared to exhibit the lowest DRC and had the highest risk of lung cancer, suggesting that a poor inducible repair response may have contributed to the excess risk. This adaptation of DRC to smoking, if it exists, appears to be long term rather than transient, because the effect was still present in former smokers but was not stronger in those who had smoked in the 24 h before the blood was drawn.

**LUC Host Cell Reactivation Assay (HCR).** To simplify the HCR assay to accommodate population studies, Qiao *et al.* (20) have replaced *CAT* with *LUC* using the plasmid expression vector *pCMVluc* (4.683 kb). This plasmid is the same construct containing a human cytomegalovirus immediate promoter and enhancer (4) except for the reporter gene. The procedures for these assays are basically the same in terms of cell culture, cell harvesting, and transfection (4). The cell extraction procedure is much simplified for the LUC assay, in which the cell pellets are frozen and thawed only once (instead of three times as for the CAT assay) in ethanol and dry ice, and 37°C water baths. For each determination of the LUC assay, LUC absorbance is measured by a luminometer. It takes only 3 h to complete the laboratory procedures for the LUC assay compared with 9 h for the CAT assay. The LUC absorbances are recorded for the cells with undamaged (background reading) and UV-damaged (repair reading) plasmids, respectively.

DRC (%) is calculated as a ratio of the damaged plasmid values to the undamaged plasmid values multiplied by 100%. The dose-repair curve for the LUC activity was first established by performing the LUC assay with three lymphoblastoid cell lines with different levels of DRC and primary lymphocytes from three donors. Assays repeated four times with one repair-proficient cell line and one repair-deficient cell line indicated that the LUC assay could easily distinguish repair-proficient from repair-deficient cell lines with relatively small coefficients of variation (5.4% and 7.2%, respectively). The DRC phenotype by the two independent CAT and LUC assays in parallel showed that they were highly correlated, with a correlation coefficient of 0.651 ($P < 0.0001$; Ref. 20), suggesting that these two assays are comparable, as we had predicted.

**Mutagen Sensitivity Assays.** Cytogenetic assays in PBLs have been extensively used to survey exposure and response of humans to genotoxic agents, based on the unproven hypothesis that the extent of genetic damage in the lymphocytes may reflect critical events in carcinogenesis in the affected tissues. The European Cohort Studies (21, 22) followed >3000 occupationally exposed individuals and documented that chromosomal damage at entry into the cohort was predictive of subsequent cancer risk. A phenotypic assay of intrinsic cancer susceptibility, the mutagen sensitivity assay, was developed by Hsu *et al.* (23) using an *in vitro* mutagen challenge to demonstrate interindividual differences in susceptibility to carcinogenic agents. In this assay, the frequency of *in vitro* bleomycin-induced breaks in short-term PBL cultures is used as a measure of cancer susceptibility (23). Hsu (24) hypothesized that in the general population, susceptibility to chromosome damage in response to clastogens varies along a continuum, with recognized chromosome fragility syndromes being at one extreme of the spectrum.

Bleomycin, one of the mutagen challenges used, is a clastogenic agent that mimics the effects of radiation by generating free oxygen radicals capable of producing DNA single- and double-strand breaks after formation of a complex with DNA, ferrous ions ($Fe^{2+}$), and oxygen, resulting in the release of oxygen radicals (25). Most of the breaks are repaired by BER. Repair is rapid, with a half-life of only a few minutes (26).

The assay requires only 1 ml of a fresh whole blood sample that undergoes short-term culture. Stimulated cells are exposed to a "pulse" of carcinogen, and after a period of time to enable repair to occur, the cells are arrested in mitosis, harvested, and stained with Giemsa. For each sample, the chromosome breaks in 50 metaphases are counted and expressed as the mean number of b/c. Chromatid gaps or attenuated regions are disregarded; only frank chromatid breaks or exchanges are recorded.

We have shown, in several smaller case-control studies, that mutagen sensitivity is an independent risk factor for lung cancer with a dose-response relationship to the numbers of induced chromosome breaks (27–30). In recently updated, but previously unpublished work, we have generated bleomycin sensitivity data on 612 lung cancer cases and 557 controls included in the DRC data analysis described above. The mean (SD) for bleomycin-induced breaks was 0.78 (0.4) b/c for the cases and 0.62 (0.33) b/c for the controls ($P < 0.001$). When dichotomized at the median control bleomycin-induced break value of 0.56 b/c, higher numbers of induced breaks were

*Table 3* Risk estimates for lung cancer associated with bleomycin sensitivity

| Bleomycin sensitivity (b/c) | Cases | Controls | Adjusted OR (95% CI) |
|---|---|---|---|
| Overall[a] | | | |
| ≤0.56 | 197 | 264 | Ref. |
| >0.56 | 415 | 293 | 1.89 (1.5–2.4) |
| Quartiles[a] | | | |
| ≤0.38 | 92 | 144 | Ref. |
| 0.38–0.56 | 105 | 120 | 1.36 (0.9–2.0) |
| 0.56–0.80 | 161 | 151 | 1.63 (1.2–2.3) |
| >0.80 | 254 | 142 | 2.81 (2.0–3.9) |
| | | | $P < 0.001$(trend) |
| Smoking Status[b] | | | |
| Never | | | |
| ≤0.56 | 22 | 29 | |
| >0.56 | 44 | 39 | 1.54 (0.7–3.2) |
| Former | | | |
| ≤0.56 | 88 | 135 | |
| >0.56 | 202 | 162 | 1.78 (1.3–2.5) |
| Current | | | |
| ≤0.56 | 87 | 100 | |
| >0.56 | 169 | 92 | 2.11 (1.4–3.1) |
| Pack-years[b] | | | |
| <39.3 | | | |
| ≤0.56 | 64 | 110 | |
| >0.56 | 122 | 121 | 1.69 (1.1–2.5) |
| ≥39.3 & <49.75 | | | |
| ≤0.56 | 30 | 33 | |
| >0.56 | 60 | 45 | 1.49 (0.8–2.8) |
| ≥49.75 | | | |
| ≤0.56 | 81 | 92 | |
| >0.56 | 189 | 88 | 2.53 (1.7–3.8) |

[a] Adjusted by age, gender, and smoking status.
[b] Adjusted by age and gender.

*Table 4* Risk estimates for lung cancer associated with BPDE sensitivity

| BPDE (b/c) | Cases | Controls | Adjusted OR (95% CI) |
|---|---|---|---|
| Overall[a] | | | |
| ≤0.54 | 158 | 200 | Ref. |
| >0.54 | 248 | 211 | 1.54 (1.2–2.0) |
| Quartiles[a] | | | |
| ≤0.36 | 62 | 99 | Ref. |
| 0.36–0.54 | 96 | 101 | 1.53 (1.0–2.3) |
| 0.54–0.80 | 119 | 105 | 1.83 (1.2–2.8) |
| >0.80 | 129 | 106 | 2.08 (1.4–3.2) |
| | | | $P < 0.001$(trend) |
| Smoking Status[b] | | | |
| Never | | | |
| ≤0.54 | 10 | 13 | |
| >0.54 | 22 | 30 | 0.90 (0.3–2.5) |
| Former | | | |
| ≤0.54 | 74 | 101 | |
| >0.54 | 112 | 118 | 1.31 (0.9–2.0) |
| Current | | | |
| ≤0.54 | 74 | 86 | |
| >0.54 | 114 | 63 | 2.12 (1.4–3.3) |
| Pack-years[b] | | | |
| <39.3 | | | |
| ≤0.54 | 48 | 81 | |
| >0.54 | 74 | 81 | 1.47 (0.9–2.4) |
| ≥39.3 & <49.75 | | | |
| ≤0.54 | 25 | 23 | |
| >0.54 | 39 | 29 | 1.28 (0.6–2.7) |
| ≥49.75 | | | |
| ≤0.54 | 75 | 83 | |
| >0.54 | 113 | 71 | 1.77 (1.2–2.7) |

[a] Adjusted by age, gender, and smoking status.
[b] Adjusted by age and gender.

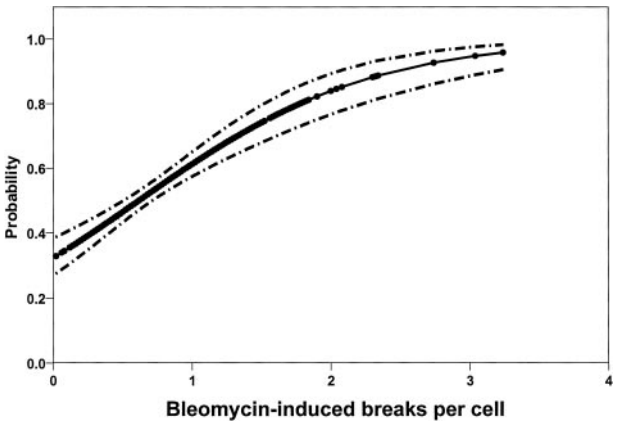


*Fig. 2.* Probability of lung cancer: univariate model for bleomycin sensitivity with 95% CIs.

associated with a univariate risk estimate of 1.89 (1.5–2.4; Table 3). As we demonstrated with DRC data, there was a similar pattern of increasing risk associated with increased numbers of induced b/c with adjusted ORs for the second, third, and fourth quartiles of 1.36 (0.9–2.0), 1.63 (1.2–2.3), and 2.81 (2.0–3.9), respectively ($P$ for trend <0.001; Table 3). The risks associated with mutagen sensitivity stratified by smoking status were 1.54 (0.7–3.2) for never smokers, 1.78 (1.3–2.5) for former smokers, and 2.11 (1.4–3.1) for current smokers. Highest risk (OR = 2.53) was evident in the heaviest pack-year tertile. Fig. 2 depicts the increasing probability of being a case (univariate model) as the level of induced damage increases.

**BPDE Sensitivity Assay.** For studying tobacco-related cancers, a more appropriate mutagen to trigger DNA damage is BPDE, the product of bioactivation *in vivo* by cytochrome P450 and epoxide hydrolase (31–33). As mentioned above, the NER pathway is responsible for removal of such bulky BPDE adducts and for the restoration of normal DNA structure (34). However, *in vitro* experiments suggest that base-excision repair may also play an important role in repair of these lesions.

We have BPDE-induced b/c data on a smaller subset of 406 cases and 411 controls. The mean (SD) b/c values were 0.69 (0.37) b/c for the cases and 0.63 (0.39) b/c for the controls ($P < 0.001$). The adjusted univariate risk estimate was 1.54 (1.2–2.0; Table 4). The risk in the highest quartile compared with the lowest was 2.08 (1.4–3.2). The adjusted risk for never smokers was 0.90 (0.3–2.5), compared with 1.31 (0.9–2.0) for former smokers and 2.12 (1.4–3.3) for current smokers. The risk in the highest pack-year stratum was 1.77 (1.2–2.7). These data additionally confirm our previously published data that BPDE-induced sensitivity is a risk factor for lung cancer (31).

The risk estimate for the combination of the two assays (suboptimal DRC and increased bleomycin sensitivity) was 2.91 (2.1–4.1) compared with 2–2.5-fold elevated risks for each assay alone (Table 5). For the combination of suboptimal DRC and BPDE sensitivity, the risk was 2.87 (1.9–4.3). Fig. 3 summarizes the probability of lung cancer by the various combinations of assay results. The curves are parallel with individuals who exhibited both adverse phenotypes having the highest probabilities followed by those who were only bleomycin sen-

*Table 5* Joint effects of DNA repair capacity and bleomcyin sensitivity

| | DRC (%) | Cases | Controls | Adjusted OR[a] (95% CI) |
|---|---|---|---|---|
| Bleomycin (b/c) | | | | |
| ≤0.56 | >8.06 | 90 | 151 | Ref. |
| ≤0.56 | ≤8.06 | 117 | 113 | 2.03 (1.4–3.0) |
| >0.56 | >8.06 | 176 | 134 | 2.45 (1.7–3.5) |
| >0.56 | ≤8.06 | 239 | 159 | 2.91 (2.1–4.1) |
| | | | | ($P < 0.001$) |
| BPDE (b/c) | | | | |
| ≤0.54 | >8.06 | 62 | 115 | Ref. |
| ≤0.54 | ≤8.06 | 96 | 85 | 2.19 (1.4–3.4) |
| >0.54 | >8.06 | 98 | 106 | 1.99 (1.2–2.7) |
| >0.54 | ≤8.06 | 150 | 105 | 2.87 (1.9–4.3) |
| | | | | ($P < 0.001$) |

[a] Adjusted by age, gender, and smoking status.

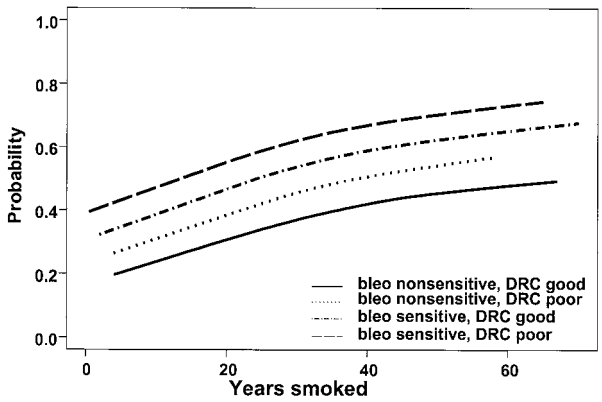


*Fig. 3.* Probability of lung cancer: multivariate model for combined bleomycin sensitivity and DRC.

sitive. There was no evidence of statistical interaction between these two assays in risk prediction.

**Comet Assay.** Over the last decade, the comet assay or single cell gel electrophoresis, has been developed as a rapid visual method for measuring DNA breakage in single cells. With advances in automated imaging technology, this assay could prove to be a promising marker to gauge host susceptibility for cancer in large molecular epidemiology studies. The comet assay appears to have many advantages including allowing relatively high-throughput screening, requiring a few cells, and facilitating the detection of primary DNA damage in individual cells (35). Although it requires viable cells, it does not require growth, and is applicable to any cell line or tissue from which a single cell suspension can be obtained, and can even be applied to terminally differentiated cells (36). We have slightly modified the comet assay under alkaline conditions as described originally by Singh *et al.* (37), and have described the methods in detail previously (38). For this assay, a cell culture is mixed with agarose gel and attached to a microscope slide. The cells are lysed by submersing the slides in freshly prepared lysis buffer for ~1 h at 4°C to remove all of the cellular proteins. To allow for DNA denaturation, unwinding, and expression of the alkali-labile sites, the slides are next placed in alkali buffer (pH 12.0). To separate the damaged DNA from the intact nuclei, a constant electric current is applied, and the slides are neutralized, fixed in 100% methanol, and stored in the dark at room temperature until ready for analysis. A fluorescent dye that binds to the DNA is used for quantification of DNA damage. Under different biochemical conditions for both lysis and electrophoresis, different classes of DNA lesions can be detected.

During electrophoresis, damaged DNA migrates from the nucleus toward the anode as a result of a constant electric current that forms the typical "comet" cell. Fifty consecutive cells (25 cells from each end of the slide) are manually selected, and comet cells are automatically quantified via the Komet 4.0.2 (Kinetic Imaging Ltd.) imaging software attached to a fluorescent microscope. The Olive tail moment parameter, [(tail mean − head mean) × (tail % DNA/100)], is used as the parameter for DNA damage and it is automatically quantified by the imaging system software (39). The "head" of the comet relates to the nucleus of the cell, whereas the "tail" of the comet refers to the damaged DNA that has been liberated from the nucleus via electrophoresis. The "tail mean − head mean" calculates the difference in the distance between the center of gravity of the DNA distribution in the comet head from the center of gravity of DNA distribution in the comet tail. For each subject, the averages of the Olive tail moment are calculated for the baseline comets and for mutagen-induced comets.

We have measured baseline comets, and BPDE- and γ-radiation induced comets in the PBLs of 162 lung cancer patients and 154 controls.[5] The cases exhibited significantly higher levels of induced DNA damage after both BPDE and γ-radiation exposure. The mean tail moment after BPDE exposure was 4.45 (1.71) for the cases compared with 3.23 (1.59) for the controls ($P < 0.01$). The comparable values for γ-radiation-induced damage were 5.59 (1.81) and 4.68 (1.69), respectively ($P < 0.01$). We have published data recently on 114 bladder cancer patients and 145 controls that show very similar patterns (38). When the tail moment was dichotomized at the 50th percentile value in the control subjects, the OR was 1.84 (95% CI, 1.1–3.2) for the baseline comet level after adjusting for age, gender, ethnicity, and smoking status; 1.81 (95% CI, 1.1–3.1) for γ-radiation-induced comet levels, and 1.69 (95% CI, 1.0–2.9) for BPDE-induced comets. A statistically significant 2-fold increased risk (OR, 2.22; 95% CI, 1.1–4.4) was evident in individuals who exhibited greater DNA damage for both mutagens. When the tail moment was categorized into tertile values, there was a positive association between increasing risk with higher levels of DNA damage across the tertiles for both the baseline and γ-radiation-induced comets.

The comet assay can also be adapted to measure DNA repair. Postexposure decrease of migration indicates that the DNA lesions are repaired. Singh *et al.* (40) and Tice (41) have shown that ~50% of the strand breaks in lymphocytes are repaired within 10–15 min, and all have been repaired within 2 h. In 100 lung cancer patients and 110 controls, Schmezer *et al.* (42) showed that cases were significantly more sensitive to bleomycin and showed reduced DRC. They measured DRC after a fixed time as the percentage of undamaged cells after bleomycin treatment divided by the percentage of undamaged cells in the control population. DRC was 68% in the cases and 81% in the controls ($P < 0.001$). Rajee-Behbahani (43) reported a DRC of 67% in 160 lung cancer cases and 79.3% in 180 controls. When the data were dichotomized at the control median of the bleomycin-induced tail moment, only 18% of the cases were below the median, and 82% were in the hypersensitive range.

The comet assay can be applied to any eukaryotic cell

[5] Unpublished observation Wu, personal communication.

population, data are obtained within a few hours of sampling, and the assay is cost effective (44). Moreover, the comet assay has the potential for clinical application. Schmezer *et al.* (42) have shown that cryopreservation for up to 12 months does not affect the sensitivity of the lymphocytes, and the reproducibility of the assay was good. Therefore, the comet assay is an applicable assay to assess *in vitro* genetic instability in large-scale epidemiological studies.

**Genetic Influence on Phenotype Assays.** Because the mutagen sensitivity and host-cell reactivation assays are labor intensive and not amenable to large-scale, high-throughput application in population studies, our long-term goal is to identify genotypes that predict the phenotypes of suboptimal DRC and mutagen sensitivity. Mohrenweiser and Jones (45) have pointed out several lines of evidence that differences in DRC reflect genetic differences. DRC in lymphocyte subpopulations, from an individual, exhibit similar repair capacities. Furthermore, intraindividual variation in repair capacity in different subpopulations of lymphocytes is significantly smaller than is interindividual variation (46). There are also published studies showing that reduced repair capacity may aggregate in first-degree relatives of cancer probands (47, 48). Ankathil *et al.* (49) showed that relatives of colorectal cancer patients tended to exhibit higher bleomycin-induced b/c. Finally, we have shown previously that the phenotype of reduced repair capacity in the NER pathway is independent of the phenotype for double-strand break repair (50).

**DNA Repair Gene Polymorphisms.** Extensive work is under way to resequence DNA repair genes to identify variations that may be associated with reduced function of their encoded proteins, rather than absence of function. Such polymorphisms could explain inter-individual differences in DRC (51–53). Although the variant alleles are likely to be associated with only modest risk, because they exist at polymorphic frequency, the attributable risks assume substantial relevance. As Berwick and Vineis (3) point out, studies that compare genetic polymorphisms with functional assays will likely be the most valuable type of study to clarify the role of a defect in DRC with the development of cancer. In the following section, we will focus on only two genes from different repair pathways.

XPD (originally named excision repair cross complementing group 2) is one of the seven genetic complementation groups encoding for proteins involved in the NER pathway. The XPD protein has a dual function in NER and in basal transcription. It functions as an evolutionary conserved ATP-dependent helicase within the multisubunit transcription repair factor complex, TFIIH. The *XPD* gene is highly conserved in eukaryotes, with homology to Rad3 and Rad15. The two polymorphisms of interest are *Lys751Gln* (exon 23) and *Asp312Asn* (exon 10; Ref. 54). The amino acid substitution *Lys751Gln* in exon 23 does not reside in a known helicase/ATPase domain, but is at an amino acid residue identical in human, mouse, hamster, and fish XPD, suggesting functional relevance for such a highly evolutionary conserved sequence (54).

Because many different mutations have been identified in the *XPD* gene, TFIIH transcriptional activity is probably relatively tolerant to amino acid changes in the XPD protein. It is also possible that mutations could destroy or alter repair function without affecting transcriptional activity. As Lunn *et al.* (55) suggested, the Lys allele may have different effects in different repair pathways assessed by different repair assays. The overall effect of conservative mutations in XPD may be subtle, because they would not alter XPB and XPD helicase activity, and multiple alterations might be needed before any effect was noted. We have published data previously on *XPD* genotype frequencies in 341 white lung cancer cases and 360 controls (52). The variant allele frequencies at codons 751 and 312 for the cases were 0.36 and 0.29, respectively, compared with 0.33 and 0.27, respectively, for the controls. The allele frequencies for the cases and controls were in Hardy-Weinberg equilibrium by the goodness-of-fit $\chi^2$ statistic. On logistic regression analysis, the adjusted ORs for the variant *Lys751Gln* or *Asp312Asn* genotypes were 1.36 and 1.51, respectively, although neither estimate was statistically significant. For individuals homozygous for the variant genotype at either locus, the adjusted risk estimate was 1.84 (1.11–3.04; $P = 0.018$ for trend).

Zhou *et al.* (56) have published data on *XPD* polymorphism frequencies in 1092 lung cancer cases, and 1240 spouse and friend controls. They reported an overall adjusted OR of 1.47 (95% CI, 1.1–2.0) for the *Asp312Asn* polymorphism (Asn/Asn *versus* Asp/Asp) and 1.06 (95% CI, 0.8–1.4) for the *Lys751Gln* polymorphism (Gln/Gln *versus* Lys/Lys). Gene-smoking interaction analyses revealed that the adjusted ORs for each of the two polymorphisms decreased significantly as pack-years increased. A stronger gene-smoking interaction was observed for the *Asp312Asn* polymorphism than for the *Lys751Gln* polymorphism. They concluded that cumulative cigarette smoking modifies the association between *XPD* polymorphisms and lung cancer risk.

Similarly, Hou *et al.* (57) reported that in 185 Swedish lung cancer cases and 162 matched population controls, the presence of one or two variant alleles was associated with increased risk for lung cancer among never-smokers only, in particular, younger (<70 years) never-smokers [OR, 2.6 (1.1–6.5) for exon 10; and OR, 3.2 (1.3–8.0) for exon 23]. The opposite findings were reported for lung cancer risk in a Chinese population (58). The *XPD* 751 Lys allele (combined Lys/Lys and Lys/Gln genotypes) was associated with a significantly increased risk of lung cancer (OR, 3.19; 95% CI, 1.01–10.07).

**Genotype Phenotype Correlations.** Amino acid differences (especially at conserved sites) in these enzymes could result in changes in repair proficiency. The next logical step is the challenging task of evaluating the functional relevance of these polymorphisms. There are a variety of factors that modulate the path from genotype to phenotype including protein-protein interactions, post-translational modification, gene silencing, epigenetic regulation, and environmental factors. Furthermore, proteins involved in DNA repair pathways are often multifunctional, resulting in a variety of phenotypes. We have evaluated the correlation between some of the DNA repair gene polymorphisms and our functional DNA repair data.

As reported previously, we noted that the *XPD* polymorphisms appeared to modulate DRC (52). Specifically, in our lung cancer cases, DRC was 8.21% for those with the Lys/Lys 751 common genotype, marginally higher (7.65%; $P = 0.10$) than in Lys/Gln heterozygotes, and significantly higher than the DRC of Gln/Gln homozygotes (7.20%; $P = 0.041$). The $P$ for the trend was 0.017. A similar trend ($P = 0.008$) was evident for DRC among the cases for the *Asp312Asn* genotypes (8.37%, 7.50%, and 6.84%, for wild-type homozygotes, heterozygotes, and variant homozygotes, respectively). These patterns were less evident among the 360 controls, although wild-type homozygous and heterozygous controls exhibited the most proficient DRC. For cases, the risk (95% CI) for exhibiting suboptimal DRC was 1.57 (0.74–3.35) for those with the 751Gln/Gln genotype, and for the variant codon 312 Asn/Asn genotype, the risk was 3.50 (1.06–11.59). For cases with at least one ho-

mozygous mutant genotype, the risk was 2.29 (1.03–5.12; $P = 0.048$ for trend). Again, there were no significant patterns among the controls, although there was a nonsignificant 41% increase in risk of suboptimal DRC for controls who were homozygous at either locus. Thus, in our dataset, these two *XPD* polymorphisms were associated consistently with lower DRC in cases with a statistically significant trend and in controls with a nonstatistically significant trend. In other words, the results suggest that these two *XPD* polymorphisms had a dominant effect on DRC in cases and a smaller effect on DRC in controls.

Recently, we confirmed the finding that the homozygous mutant genotypes were associated with less efficient DRC, in a different study population and using a different mutagen challenge, UV 800 J/m$^2$ exposure, that like BPDE invokes NER (53). Specifically, lower DRC, as measured by the LUC assay in 102 healthy non-Hispanic white subjects, was consistently associated with homozygosity of the *XPD* polymorphic alleles, representing a reduction of >10% in DRC compared with their wild-types. For the *Gln751Gln* genotype compared with the *Lys751Lys* genotype this difference was statistically significant ($P = 0.004$; Ref. 53).

Furthermore, we recently published correlative data on the comet assay and *XPD* genotypes in the previously cited bladder cancer case control study (38). Among the control subjects, nonsignificantly lower levels of BPDE-induced comets were noted for the wild-type genotypes of *XPD* exon 10 (Asp312Asn; $n = 65$; $3.23 \pm 1.16$) as compared with the variant genotypes ($n = 75$; $3.46 \pm 1.09$; $P = 0.236$). The variant *XPD* exon 23 polymorphism (Lys751Gln) genotypes were also associated with greater numbers of BPDE-induced comets ($n = 79$; $3.46 \pm 1.10$) as compared with the wild-type genotypes ($n = 62$; $3.24 \pm 1.15$), but this difference was not statistically significant ($P = 0.247$). Likewise, similar patterns were noted for BPDE-induced DNA damage associated with the variant genotypes of *XPD* exons 10 and 23 ($n = 64$; $3.45 \pm 1.12$) as compared with the wild-type genotypes of *XPD* exons 10 and 23 ($n = 50$; $3.13 \pm 1.14$). However, this difference was also not statistically significant ($P = 0.135$).

These data are consistent with some of the published small-scale studies looking at such genotype-phenotype correlations. Hou *et al.* (57) noted a significant trend for increasing aromatic DNA adduct levels with increasing number of variant alleles in exon 10 ($P = 0.02$) or in exon 23 ($P = 0.001$). In addition, subjects with the combined exon 10 AA and exon 23 CC genotype showed a significantly higher levels compared with those with any of the other genotypes ($P = 0.02$). Hu *et al.* (59) reported that in both prostate cancer cases ($n = 66$) and controls ($n = 54$), those homozygous for the variant allele had lower DRC (8.7% and 6.4%, respectively) than those with the wild-type genotype (11.1% and 10.9%).

However, Moller *et al.* (60) reported no relationship of the *Lys751Gln* polymorphism with DRC (measured by the host cell reactivation assay or the comet assay) in 80 subjects, including 20 healthy subjects. Lunn *et al.* (55) reported that possessing the Lys/Lys 751 common *XPD* genotype was associated with increased risk of suboptimal DRC (as reflected in the number of X-ray-induced lymphocyte chromatid aberrations). They found no association with the Asn312 allele. In another small sample of 76 healthy subjects, no association was noted between SCE frequencies or in the presence of polyphenol DNA adducts by *Lys751Gln* genotype (61).

For a complex disease like cancer, multiple genes, each with a small effect, probably act independently or interact with other genes to influence the disease phenotype. Although these data suggest that the polymorphisms have functional relevance, biochemical and biological characterization of the variants are needed to validate the findings.

**XRCC1.** The X-ray repair cross-complementing group 1 gene (*XRCC1*) encodes a protein that functions in the repair of single-strand breaks, the most common lesions in cellular DNA, through one of the most highly conserved DNA repair mechanisms, BER (62). BER involves excision of the damaged region followed by repair synthesis using the opposite strand as the template. The XRCC1 protein complexes with DNA ligase III, DNA polymerase $\beta$, and poly(ADP-ribose) polymerase to rejoin DNA strand breaks and repair gaps left during BER (63–65). A polymorphism at *XRCC1* 28152 site (G→A) of codon 399 in exon 10 results in a nonconservative amino acid substitution of arginine for glutamine. The polymorphism at *XRCC1* 26304 site (C→T) of codon 194 in exon 6 results in a nonconservative amino acid substitution of arginine for tryptophan (54).

In a recent evaluation of 524 lung cancer cases and 524 controls, we found that the allele frequencies were 0.07 and 0.38 for *XRCC1* exon 6 and exon 10 variant alleles, respectively, which are consistent with many other studies (66, 67). The differences of allele frequencies detected among these studies might be because of different sample sizes and study populations.

We used bleomycin-induced b/c data from 524 lung cancer cases and 524 controls to correlate with genotype data, because bleomycin induces strand breaks requiring BER (Table 6). Individuals with the wild-type (*Trp1194 Trp*) genotype exhibited higher numbers of bleomycin-induced b/c than those with the combined variant genotype ($P = 0.007$ and 0.005 for cases and controls, respectively). Similarly, among both the cases and the controls, individuals with the homozygous variant genotype at exon 10 exhibited higher b/c than those with the wild-type genotype, with evidence of a gene dosage effect, although the trend was not statistically significant. Controls who were homozygous variant at this locus exhibited a mean score of $0.71 \pm 0.37$ that approximated the scores evident among the cases, and that was significantly different from the scores for the wild-type genotype ($P = 0.03$). When we combined the two genotypes, and used the exon 6 *Arg/Trp* and *Trp/Trp* combined genotype and exon 10 *Arg 399 Arg* genotype as the reference category, these differences were enhanced ($P$ for trend = 0.049 for the cases and 0.032 for the controls). Among the cases, the scores were 0.80 and 0.84 for one adverse allele and >1 respectively, compared with 0.63 and 0.72 for the controls. These patterns remained consistent when the data were stratified by gender, age, and smoking status.

These findings are biologically plausible. The DNA repair-deficient cell line, EM-11, which has high background SCE frequencies, has a mutation in XRCC1 at codon 390 (54). This cell line is also sensitive to accumulation of single-strand DNA breaks after damage. The *XRCC1* codon 399 is located within the BRCT domain (amino acids 301–402) that interacts with poly(ADR-ribose), and that is present in many DNA damage response and cell-cycle checkpoint proteins. This region also has homology with yeast Rad4 repair-related genes. Because the role of *XRCC1* in BER brings together DNA polymerase $\beta$, DNA ligase III, and poly(ADP-ribose) polymerase at the site of DNA damage, the exon10 variant could have an altered repair activity. The codon 194 polymorphism is in the linker region of the XRCC1 $NH_2$-terminal domain separating the helix 3 and $\beta$-pol involved in binding a single-nucleotide gap DNA substrate (68). Therefore, it is less likely

*Table 6* Distribution of bleomycin sensitivity (b/c) by XRCC1 genotype in lung cancer cases and controls

| Genotype | Cases | | | | Controls | | | |
|---|---|---|---|---|---|---|---|---|
| | *n* | Mean | SD | *P* | *n* | Mean | SD | *P* |
| Exon 6 | | | | | | | | |
| Arg 1194 Arg, Arg 1194 Trp | 72 | 0.70 | 0.32 | Ref. | 73 | 0.56 | 0.24 | Ref. |
| Trp 1194 Trp | 452 | 0.81 | 0.42 | 0.007 | 451 | 0.65 | 0.35 | 0.005 |
| Exon 10 | | | | | | | | |
| Arg 399 Arg | 218 | 0.77 | 0.41 | Ref. | 208 | 0.61 | 0.33 | Ref. |
| Arg 399 Gln | 240 | 0.81 | 0.40 | 0.373 | 237 | 0.64 | 0.33 | 0.417 |
| Gln 399 Gln | 66 | 0.84 | 0.48 | 0.288 | 79 | 0.71 | 0.37 | 0.033 |
| | | *P*(trend) = 0.225 | | | | *P*(trend) = 0.040 | | |
| Exons 6 and 10 | | | | | | | | |
| 6–Arg/Arg, Arg/Trp | | | | | | | | |
| 10–Arg/Arg | 41 | 0.67 | 0.32 | Ref. | 43 | 0.60 | 0.24 | Ref. |
| 10–Arg/Gln | 418 | 0.80 | 0.41 | 0.013 | 406 | 0.63 | 0.33 | 0.509 |
| 6–Trp/Trp | | | | | | | | |
| 10–Gln/Gln | 65 | 0.84 | 0.48 | 0.025 | 75 | 0.72 | 0.37 | 0.038 |
| | | *P*(trend) = 0.049 | | | | *P*(trend) = 0.032 | | |

to cause a significant change in repair function. In fact, our data showed that individuals with the variant allele had fewer b/c than those who had the wild-type genotype.

There have been a few previous studies evaluating the functional significance of these polymorphisms. Lunn *et al.* (66) noted that individuals with the *399 Gln* allele were at significantly higher risk (OR, 2.4) for exhibiting detectable aflatoxin B1 adducts and higher glycophorin A variant frequency than 399Arg/Arg carriers. These data suggest that the 399 Gln allele is involved both in BER and recombination repair. They also reported a gene-dosage effect (66). However, no significant effects were noted for other *XRCC1* polymorphisms. Duell *et al.* (61) found elevated SCE frequencies and polyphenol adducts with *399 Gln/Gln* homozygous genotypes. Abdel-Rahman and el-Zein (69) evaluated functionality of the 399 Gln polymorphism in 47 volunteers by measuring 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone-induced SCEs. They noted that individuals carrying the Gln allele had significantly higher numbers of SCEs in response to 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone treatment compared with those with the Arg/Arg genotype. No differences were detected with the codon 194 genotype.

Matullo *et al.* (70) evaluated the levels of WBC DNA adducts in 308 healthy Italian individuals and found higher adduct levels among never smokers who were 399 Gln homozygous (15.6 *versus* 6.78 in those who had the wild-type genotype; $P = 0.007$). We also observed a gene-dosage effect for the 399 Gln allele. Whereas studies of actual protein function and related DNA repair phenotype are needed to confirm that these amino acid differences do modulate DRC, these data are indeed suggestive, and provide additional confirmation that these polymorphisms may be important in cancer risk assessment.

**The Missing Pieces of the Puzzle.** We now have an extensive set of data on phenotypic and genotypic markers of susceptibility as predictors of lung cancer risk. However, there are many issues that need to be addressed. These issues include selection bias, variability and reliability of the tests, and the retrospective nature of the study designs, raising concerns about the impact of disease status on the assay results. We also need to know how well these functional data (derived from surrogate lymphocytic tissue) reflect events at the level of the target tissue. Therefore, the logical next step in risk assessment is to correlate these surrogate phenotype/genotype data with tissue-specific analyses from bronchial biopsies, bronchial washings, and tumor tissue.

There is a paucity of data on such analyses. Kleinsasser *et al.* (71) reported a poor correlation between genotoxic sensitivity of PBLs and mucosal cells of the upper aerodigestive tract. However, Auckley *et al.* (72) have shown a very tight correlation coefficient of 0.83 ($P = 0.003$) between DNA-dependent protein kinase activity that is activated by the presence of double-strand breaks in peripheral mononuclear cells and in bronchial epithelial cells, obtained at bronchoscopy from 10 individuals. They also repeated the experiments after *in vitro* 5-$\mu$g bleomycin exposure and noted a correlation coefficient of 0.59 ($P = 0.03$) in 13 subjects. The association was less strong at higher doses of bleomycin suggesting an increase in cell killing at higher doses. Wu *et al.* (73) have shown previously that after exposure to bleomycin, the proportion of chromosome 5 aberrations is significantly higher in surviving lymphocyte from lung cancer cases than controls. Sensitivity to bleomycin, which has been correlated with lung cancer risk (73), appears linked to the survival of cells harboring chromosomal aberrations. These data support the notion that the use of peripheral lymphocytes may be an appropriate surrogate cell type for future population-based studies.

**"Integrative Epidemiology."** As the need grows for rapid and efficient translation of emerging new technologies and approaches to improved patient management, early detection, and prevention, there will inevitably be a growing emphasis placed on molecular epidemiology research and the application of the approaches intrinsic to epidemiology to other aspects of translational research. Molecular epidemiologists will increasingly be called on to identify high-risk (susceptible) subgroups who might benefit disproportionately from screening or chemoprevention interventions. Additional studies in these high-risk individuals can also provide insight into applying these approaches to the average risk population. Surrogate markers found to be most representative of molecular changes in the target tissue will be used to develop quantitative risk assessment models. Furthermore, these markers are of potential value as an adjunct to new screening modalities, such as spiral computed tomography of the lung. These approaches can also be applied to other readily accessible tissues or fluids such as urine, stool, or sputum, that provide an opportunity for relatively noninvasive evaluation of risk, and of physiological and pathophysiological states.

As it becomes increasingly important to identify cancer patient subgroups with a good prognosis or likelihood to respond to particular therapies, this approach can be extended to the study of markers of prognosis and prediction of therapeutic response. Pharmacogenetic profiles could be used to individualize therapy and to understand the functional consequences of chemoprevention, chemotherapy, or radiotherapy response.

Identification of protein patterns in serum using high-throughput proteomics linked to novel bioinformatics approaches can be useful for studying the entire spectrum from predisposition to early diagnosis and cancer outcome. Tumor DNA can be isolated from serum or plasma, as a useful source for screening specific transcripts or mutations in mitochondrial or nuclear DNA sequences, and that may have a potential role in early detection. Linking of genome-wide polymorphism analysis, DNA copy number, epigenetic changes, transcriptional profiling, and proteomics provides powerful new approaches that cannot be successfully accomplished by any discipline independently. Integrating these new approaches will require the rigor of data and sample collection and validation inherent in epidemiological research, and mandates linking of tissue repositories with well-characterized epidemiological, clinical, and follow-up data.

In summary, the approaches of molecular epidemiology with the ability to obtain and validate patient data and samples can provide a major integrative force for these disciplines. The culture of the research must be to support and nurture such high-risk, but high-impact, cutting-edge research, to promote data sharing, and data and tissue linkages, and to facilitate easy access for researchers to such data and tissue resources, with proper consideration of patient protection issues.

## Acknowledgments

We thank Dr. Gordon Mills for incisive scientific input, Susan Honn and her study staff for subject recruitment, and Joanne Sider for editorial assistance.

## References

1. Shopland, D. R., Eyre, H. J., and Pechacek, T. F. Smoking-attributable cancer mortality in 1991: Is lung cancer now the leading cause of death among smokers in the United States? J. Natl. Cancer Inst., *83:* 1142–1148, 1991.

2. Tlysty, T. D., Briot, A., Gualberto, A., Hall, I., Hess, S., Hixon, M., Kuppuswamy, D., Romanov, S., Sage, M., and White, A. Genomic instability and cancer. Mutat. Res., *337:* 1–7, 1995.

3. Berwick, M., and Vineis, P. Markers of DNA repair and susceptibility to cancer in humans: an epidemiologic review. J. Natl. Cancer Inst., *92:* 874–897, 2000.

4. Athas, A. F., Hedayati, M., Matanoski, G. M., Farmer, E. R., and Grossman, L. Development and field-test validation of an assay for DNA repair in circulating human lymphocytes. Cancer Res., *51:* 5786–5793, 1991.

5. Phillips, D. H. Fifty years of B(a)P. Nature (Lond.), *303:* 468–472, 1983.

6. Gelboin, H. V. Benzo[a]pyrene metabolism, activation, and carcinogenesis: role and regulation of mixed function oxidases and related enzymes. Physiol. Rev., *60:* 1107–1166, 1980.

7. Tang, M. S., Pierce, J. R., Doisy, R. P., Nazimiec, M. E., and MacLeod, M. C. Differences and similarities in the repair of two benzo[a]pyrene diol epoxide isomers induced DNA adducts by uvrA, uvrB, and uvrC gene products. Biochemistry, *31:* 8429–8436, 1992.

8. Koch, K. S., Fletcher, R. G., Grond, M. P., Inyang, A. I., Lu, X. P., Brenner, D. A., and Leffert, H. L. Inactivation of plasmid reporter gene expression by one benzo(a)pyrene diol-epoxide DNA adduct in adult rat hepatocytes. Cancer Res., *53:* 2279–2286, 1993.

9. Wei, Q., Cheng, L., Amos, C. I., Wang, L. E., Guo, Z., Hong, W. K., and Spitz, M. R. Repair of tobacco carcinogen-induced DNA adducts and lung cancer risk: A molecular epidemiological study. J. Natl. Cancer Inst., *92:* 1764–1772, 2000.

10. Barret, J. M., Calsou, P., and Salles, B. Deficient nucleotide excision repair activity in protein extracts from normal human lymphocytes. Carcinogenesis (Lond.), *16:* 1611–1616, 1995.

11. Shivji, M. K., Kenny, M. K., and Wood, R. D. Proliferating cell nuclear antigen is required for DNA excision repair. Cell, *69:* 367–374, 1992.

12. Cheng, L., Bucana, C. D., and Wei, Q. Fluorescence *in situ* hybridization method for measuring transfection efficiency. Biotechniques, *21:* 486–491, 1996.

13. Wei, Q., Matanoski, G. M., Farmer, E. R., Hedayati, M. A., and Grossman, L. DNA repair and aging in basal cell carcinoma: a molecular epidemiology study. Proc. Natl. Acad. Sci. USA, *90:* 1614–1618, 1993.

14. Wei, Q., Cheng, L., Hong, W. K., and Spitz, M. R. Reduced DNA repair capacity in lung cancer patients. Cancer Res., *56:* 4103–4107, 1996.

15. Risch, H. A., Howe, G. R., Jain, M., Burch, J. D., Holowaty, E. J., and Miller, A. B. Are female smokers at higher risk for lung cancer than male smokers? A case–control analysis by histologic type. Am. J. Epidemiol., *138:* 281–293, 1993.

16. Zang, E. A., and Wynder, E. L. Differences in lung cancer risk between men and women: examination of the evidence. J. Natl. Cancer Inst., *88:* 183–192, 1996.

17. Eller, M. S., Maeda, T., Magnoni, C., Atwal, D., and Gilchrest, B. A. Enhancement of DNA repair in human skin cells by thymidine dinucleotides: evidence for a p53-mediated mammalian SOS response. Proc. Natl. Acad. Sci. USA, *94:* 12627–12632, 1997.

18. Cheng, L., Eicher, S. A., Guo, Z., Hong, W. K., Spitz, M. R., and Wei, Q. Reduced DNA repair capacity in head and neck cancer patients. Cancer Epidemiol. Biomark. Prev., *7:* 465–468, 1998.

19. Wei, Q., Eicher, S. A., Guan, Y., Cheng, L., Xu, J., Young, L. N., Saunders, K. C., Jiang, H., Hong, W. K., Spitz, M. R., and Strom, S. S. Reduced expression of hMLH1 and hGTBP/hMSH6: a risk factor for head and neck cancer. Cancer Epidemiol. Biomark. Prev., *7:* 309–314, 1998.

20. Qiao, Y., Spitz, M. R., Guo, Z., Hedayati, M., Grossman, L., Kraemer, K. H., and Wei, Q. Rapid assessment of repair of ultraviolet DNA damage with a modified host-cell reactivation assay using a luciferase reporter gene and correlation with polymorphisms of DNA repair genes in normal human lymphocytes. Mutat. Res., *509:* 165–174, 2002.

21. Hagmar, L., Brogger, A., Hansteen, I., Heim, S., Hogstedt, B., Knudsen, L., Lambert, B., Linnainmaa, K., Mitelman, F., Nordenson, I., Reuterwall, C., Salomaa, S., Skerfving, S., and Sorsa, M. Cancer risk in humans predicted by increased levels of chromosome aberrations in lymphocytes: Nordic study group on the health risk of chromosome damage. Cancer Res., *54:* 2919–2922, 1994.

22. Bonassi, S., Abbondandolo, A., Camurri, L., Dal Pra, L., De Ferrari, M., Degrassi, F., Forni, A., Lamberti, L., Lando, C., Padovani, P., Sbrana, I., Vecchio, D., and Puntoni, R. Are chromosome aberrations in circulating lymphocytes predictive of future cancer onset in humans? Preliminary results of an Italian cohort study. Cancer Genet. Cytogenet., *79:* 133–135, 1995.

23. Hsu, T. C., Johnston, D. A., Cherry, L. M., Ramkisson, D., Schantz, S. P., Jessup, J. M., Winn, R. J., Shirley, L., and Furlong, C. Sensitivity to genotoxic effects of bleomycin in humans: possible relationship to environmental carcinogenesis. Int. J. Cancer, *43:* 403–409, 1989.

24. Hsu, T. C. Genetic instability in the human population: a working hypothesis. Hereditas, *98:* 1–9, 1983.

25. Burger, R. M., Peisach, J., and Horwitz, S. B. Mechanism of bleomycin action: *in vitro* studies. Life Sci., *28:* 715–727, 1981.

26. Lopex-Larraza, D., DeLucca, J. C., and Bianchi, N. O. The kinetics of DNA damage by bleomycin in mammalian cells. Mutat. Res., *232:* 57–61, 1990.

27. Spitz, M. R., Hsu, T. C., Wu, X. F., Fueger, J. J., Amos, C. I., Roth, J. A. Mutagen sensitivity as a biologic marker of lung cancer risk in African Americans. Cancer Epidemiol. Biomark. Prev., *4:* 99–103, 1995.

28. Strom, S. S., Wu, X., Sigurdson, A. J., Hsu, T. C., Fueger, J. J., Lopez, J., Tee, P. G., and Spitz, M. R. Lung cancer, smoking patterns, and mutagen sensitivity in Mexican-Americans. J. Natl. Cancer Inst. Monog, *18:* 29–33, 1995.

29. Wu, X., Delclos, G. L., Annegers, F. J., Bondy, M. L., Honn, S. E., Henry, B., Hsu, T. C., and Spitz, M. R. A case-control study of wood-dust exposure, mutagen sensitivity, and lung-cancer risk. Cancer Epidemiol. Biomark. Prev., *4:* 583–588, 1995.

30. Wu, X., Gu, J., Amos, C. I., Jiang, H., Hong, W. K., and Spitz, M. R. A parallel study of *in vitro* sensitivity to benzo[a]pyrene diol epoxide and bleomycin in lung cancer cases and controls. Cancer (Phila.), *83:* 1118–27, 1998.

31. Wei, Q., Gu, J., Cheng, L., Bondy, M. L., Jiang, H., Hong, W. K., and Spitz, M. R. Benzo(a) pyrene diol epoxide-induced chromosomal aberrations and risk of lung cancer. Cancer Res., *56:* 3975–3979, 1996.

32. Wu, X., Gu, J., Hong, W. K., Lee, J. J., Amos, C. I., Jiang, H., Winn, R. J., Fu, K. K. W., Cooper, J., and Spitz, M. R. Benzo[a]pyrene diol epoxide and bleomycin sensitivity and susceptibility to cancer of upper aerodigestive tract. J. Natl. Cancer Inst., *90:* 1393–1399, 1998.

33. Wu, X., Gu, J., Patt, Y., Hassan, M., Spitz, M. R., Beasley, R. P., and Hwang, L. Y. Mutagen sensitivity as a susceptibility marker for human hepatocellular carcinoma. Cancer Epidemiol. Biomark. Prev., *7:* 567–570, 1998.

34. Sancar, A. DNA repair in humans. Annu. Rev. Genet., *29:* 69–105, 1995.

35. Tice, R. R., Agurell, E., Anderson, D., Burlinson, B., Hartmann, A., Kobayashi, H., Miyamae, Y., Rojas, E., Ryu, J. C., and Sasaki, Y. F. Single cell gel/comet assay: guidelines for *in vitro* and *in vivo* genetic toxicology testing. Environ. Mol. Mutagen., *35:* 206–221, 2000.

36. Ross, G. M., McMillan, T. J., Wilcox, P., and Collins, A. R. The single cell microgel electrophoresis assay (comet assay): technical aspects and applications. Mutat. Res., *337:* 57–60, 1995.

37. Singh, N. P., McCoy, M. T., Tice, R. R., and Schneider, E. L. A simple technique for quantitation of low levels of DNA damage in individual cells. Exp. Cell Res., *175:* 184–189, 1988.

38. Schabath, M. B., Spitz, M. R., Grossman, H. B., Zhang, K., Dinney, C. P., Zheng, P. J., and Wu, X. Genetic instability in bladder cancer assessed by the comet assay. J. Natl. Cancer Inst., *95:* 540–547, 2003.

39. Olive, P. L., Banath, J. P., and Durand, R. E. Heterogeneity in radiation-induced DNA damage and repair in tumor and normal cells measured using the "comet" assay. Radiat. Res., *122:* 86–94, 1990.

40. Singh, N. P., Danner, D. B., Tice, R. R., Brant, L., and Schneider, E. L. DNA damage and repair with age in individual human lymphocytes. Mutat. Res., *237:* 123–130, 1990.

41. Tice, R. R. The single cell gel/comet assay: a microgel electrophoretic technique for the detection of DNA damage and repair in individual cells. *In:* D. H. Phillips, and S. Venitt (eds.), Environmental Mutagenesis, pp. 315–339. OS Scientific Publishers, 1995.

42. Schmezer, P., Rajaee-Behbahani, N., Risch, A., Thiel, S., Rittgen, W., Drings, P., Dienemann, H., Kayser, K. W., Schulz, V., and Bartsch, H. Rapid screening assay for mutagen sensitivity and DNA repair capacity in human peripheral blood lymphocytes. Mutagenesis, *16:* 25–30, 2001.

43. Rajee-Behbahani, N., Schmezer, P., Risch, A., Rittgen, W., Kayser, K. W., Dienemann, H., Schulz, V., Drings, P., Thiel, S., and Bartsch, H. Altered DNA repair capacity and bleomycin sensitivity as risk markers for non-small cell lung cancer. Int. J. Cancer, *95:* 86–91, 2001.

44. Rojas, E., Lopez, M. C., and Valverde, M. Single cell gel electrophoresis assay: methodology and applications. J. Chromatogr. B Biomed. Sci. Appl., *722:* 225–254, 1999.

45. Mohrenweiser, H. W., and Jones, I. M. Variation in DNA repair is a factor in cancer susceptibility: a paradigm for the promises and perils of individual and population risk estimation? Mutat. Res., *400:* 15–24, 1998.

46. Crompton, N. E., and Ozsahin, M. A versatile and rapid assay of radiosensitivity of peripheral blood leukocytes based on DNA and surface-marker assessment of cytotoxicity. Radiat. Res., *147:* 55–60, 1997.

47. Helzlsouer, K. J., Harris, E. L., Parshad, R., Fogel, S., Bigbee, W. L., and Sanford, K. K. Familial clustering of breast cancer: possible interaction between DNA repair proficiency and radiation exposure in the development of breast cancer. Int. J. Cancer, *64:* 14–17, 1995.

48. Bondy, M. L., Sptz, M. R., Halabi, S., Fueger, J. J., Schantz, S. P., Sample, D., and Hsu, T. C. Association between family history of cancer and mutagen sensitivity in upper aerodigestive tract cancer patients. Cancer Epidemiol. Biomark. Prev., *2:* 103–106, 1993.

49. Ankathil, R., Jyothish, B., Madhavan, J., and Nair, M. K. Deficient DNA repair capacity: a predisposing factor and high risk predictive marker in familial colorectal cancer. J. Exp. Clin. Cancer Res., *18:* 33–37, 1999.

50. Wu, X., Gu, J., Hsu, T. C., Hong, W. K., Shi, H., and Spitz, M. R. Differential sensitivity to benzo(a)pyrene diol-epoxide (BPDE) as a marker of lung cancer susceptibility. Proc. Am. Assoc. Cancer Res., *38:* 618, 1997.

51. Wu, X., Zhao, H., Wei, Q., Amos, C. I., Zhang, K., Guo, Z., Qiao, Y., Hong, W. K., and Spitz, M. R. XPA polymorphism associated with reduced lung cancer risk and a modulating effect on nucleotide excision repair capacity. Carcinogenesis (Lond.), *24:* 505–509, 2003.

52. Spitz, M. R., Wu, X., Wang, Y., Wang, L. E., Shete, S., Amos, C. I., Guo, Z., Lei, L., Mohrenweiser, H., and Wei, Q. Modulation of nucleotide excision repair capacity by XPD polymorphisms in lung cancer patients. Cancer Res., *61:* 1354–1357, 2001.

53. Qiao, Y., Spitz, M. R., Shen, H., Guo, Z., Shete, S., Hedayati, M., Grossman, L., Mohrenweiser, H., and Wei, Q. Modulation of repair of ultraviolet damage in the host-cell reactivation assay by polymorphic XPC and XPD/ERCC2 genotypes. Carcinogenesis (Lond.), *23:* 295–299, 2002.

54. Shen, M. R., Jones, I. M., and Mohrenweiser, H. Nonconservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. Cancer Res., *58:* 604–608, 1998.

55. Lunn, R. M., Helzlsouer, K. J., Parshad, R., Umbach, D. M., Harris, E. L., Sanford, K. K., and Bell, D. A. XPD polymorphisms: effects on DNA repair proficiency. Carcinogenesis (Lond.), *21:* 551–555, 2000.

56. Zhou, W., Liu, G., Miller, D. P., Thurston, S. W., Xu, L. L., Wain, J. C., Lynch, T. J., Su, L., and Christiani, D. C. Gene-environment interaction for the ERCC2 polymorphisms and cumulative cigarette smoking exposure in lung cancer. Cancer Res., *62:* 1377–1381, 2002.

57. Hou, S. M., Falt, S., Angelini, S., Yang, K., Nyberg, F., Lambert, B., and Hemminki, K. The XPD variant alleles are associated with increased aromatic DNA adduct level and lung cancer risk. Carcinogenesis (Lond.), *23:* 599–603, 2002.

58. Chen, S., Tang, D., Xue, K., Xu, L., Ma, G., Hsu, Y., and Cho, S. S. DNA repair gene XRCC1 and XPD polymorphisms and risk of lung cancer in a Chinese population. Carcinogenesis (Lond.), *23:* 1321–1325, 2002.

59. Hu, J. J., Hall, C., Torti, F. M., McCullough, D. L., Lohman, K., and Grossman, L. Deficient DNA repair and human prostate cancer risk. Proc. Am. Assoc. Cancer Res., *41:* 360, 2000.

60. Moller, P., Knudsen, L. E., Frentz, G., Dybdahl, M., Wallin, H., and Nexo, B. A. Seasonal variation of DNA damage and repair in patients with non-melanoma skin cancer and referents with and without psoriasis. Mutat. Res., *407:* 25–34, 1998.

61. Duell, E. J., Wiencke, J. K., Cheng, T. J., Varkonyi, A., Zuo, Z. F., Ashok, T. D. S., Mark, E. J., Wain, J. C., Christiani, D. C., and Kelsey, K. T. Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers of DNA damage in human blood mononuclear cells. Carcinogenesis (Lond.), *21:* 965–971, 2002.

62. Eisen, J. A., and Hanawalt, P. C. A phylogenomic study of DNA repair genes, proteins, and processes. Mutat. Res., *435:* 171–213, 1999.

63. Caldecott, K. W., McKeown, C. K., Tucker, J. D., Ljungquist, S., and Thompson, L. H. An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. Mol. Cell. Biol., *14:* 68–76, 1994.

64. Kubota, Y., Nash, R. A., Klungland, A., Schar, P., Barnes, D. E., and Lindahl, T. Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase $\beta$ and the XRCC1 protein. EMBO J., *15:* 6662–6670, 1996.

65. Ding, R., and Smulson, M. Depletion of nuclear poly(ADP-ribose) polymerase by antisense RNA expression: influences on genomic stability, chromatin organization, and carcinogen cytotoxicity. Cancer Res., *54:* 4627–4634, 1994.

66. Lunn, R. M., Langlois, R. G., Hsieh, L. L., Thompson, C. L., and Bell, D. A. XRCC1 polymorphisms: effects on aflatoxin B1-DNA adducts and glycophorin A variant frequency. Cancer Res., *59:* 2557–2561, 1999.

67. Sturgis, E. M., Castillo, E. J., Li, L., Zheng, R., Eicher, S. A., Clayman, G. L., Strom, S. S., Spitz, M. R., and Wei, Q. Polymorphisms of DNA repair gene XRCC1 in squamous cell carcinoma of the head and neck. Carcinogenesis (Lond.), *20:* 2125–2129, 1999.

68. Marintchev, A., Mullen, M. A., Maciejewski, M. W., Pan, B., Gryk, M. R., and Mullen, G. P. Solution structure of the single-strand break repair protein XRCC1 N-terminal domain. Nat. Struct. Biol., *6:* 884–893, 1999.

69. Abdel-Rahman, S. Z., and El-Zein, R. A. The 399Gln polymorphism in the DNA repair gene XRCC1 modulates the genotoxic response induced in human lymphocytes by the tobacco-specific nitrosamine NNK. Cancer Lett., *159:* 63–71, 2000.

70. Matullo, G., Palli, D., Peluso, M., Guarrera, S., Carturan, S., Celentano, E., Krogh, V., Munnia, A., Tumino, R., Polidoro, S., Piazza, A., and Vineis, P. XRCC1, XRCC3, XPD gene polymorphisms, smoking and (32)P-DNA adducts in a sample of healthy subjects. Carcinogenesis (Lond.), *22:* 1437–1445, 2001.

71. Kleinsasser, N. H., Wallner, B. C., Kastenbauer, E. R., Muenzenrieder, R. K., and Harreus, U. A. Comparing the genotoxic sensitivities of human peripheral blood lymphocytes and mucosa cells of the upper aerodigestive tract using the Comet assay. Mutat. Res., *467:* 21–30, 2000.

72. Auckley, D. H., Crowell, R. E., Heaphy, E. R., Stidley, C. A., Lechner, J. F., Gilliland, F. D., and Belinsky, S. A. Reduced DNA-dependent protein kinase activity is associated with lung cancer. Carcinogenesis (Lond.), *22:* 723–727, 2001.

73. Wu, X. F., Spitz, M. R., Delclos, G. L., Connor, T. H., Zhao, Y., Siciliano, M. J., and Hsu, T. C. Survival of cells with bleomycin-induced chromosomal lesions in the cultured lymphocytes of lung cancer patients. Cancer Epidemiol. Biomark. Prev., *5:* 527–532, 1996.

# Genetic Susceptibility to Lung Cancer: The Role of DNA Damage and Repair ,

Margaret R. Spitz, Qingyi Wei, Qiong Dong, et al.

*Cancer Epidemiol Biomarkers Prev* 2003;12:689-698.