Cancer-related changes in prostate DNA as men age and early identification of metastasis in primary prostate tumors

Donald C. Malins*†, Paul M. Johnson*, Edward A. Barker‡, Nayak L. Polissar§, Thomas M. Wheeler¶, and Katie M. Anderson*

*Biochemical Oncology Program, Pacific Northwest Research Institute, 720 Broadway, Seattle, WA 98122; ‡Molecular Oncology International, Seattle, WA 98125; §Mountain-Whisper-Light Statistical Consulting, Seattle, WA 98112-2913; and ¶Department of Pathology, Baylor College of Medicine, Houston, TX 77030-3411

Contributed by Donald C. Malins, March 10, 2003

Using statistical analyses of Fourier transform-IR spectra, we show that DNA of the histologically normal prostates of men 16–80 years old undergoes structural changes in the bases and backbone with increasing age. Of the older men (ages 55–80), 42% exhibited a DNA phenotype mimicking that of primary prostate tumors from a comparable age group. This cancer-like phenotype, which was not found in the younger men (ages 16–36), appears to arise from progressive age-related damage to DNA. The mean concentrations of 8-hydroxypurine lesions (e.g., 8-hydroxyguanine) were substantially higher for the older men than for the younger men. This finding suggests that the hydroxyl radical contributed to the structural changes that characterize the cancer-like phenotype. Strikingly, we were additionally able to discriminate between the DNA of primary prostate tumors and the DNA of primary prostate tumors from which distant metastases had been identified. Moreover, logistic regression analysis was able to predict the probability that a tumor had metastasized with ≈90% sensitivity and specificity. Collectively, these findings are particularly promising for identifying men at risk for developing prostate cancer, as well as for the early determination of whether a primary tumor has progressed to the metastatic state. This is highly important because the prognosis of histologically similar prostate carcinomas varies, thus creating a need to predict which cancers are most likely metastatic.

In the United States, prostate cancer is the most prevalent of all cancers and the second-leading cause of cancer mortality in men (1). The incidence rates for this disease rise markedly as a function of increasing age (1–3). Notably, 70% of all newly diagnosed prostate cancers occur in men over age 65 (2). The greatest threat from prostate cancer lies in its potential to metastasize (3). Clinical diagnosis of metastatic prostate cancer presently depends on histological identification of distant metastases. Yet, by the time the tumor has metastasized, the success of intervention is severely limited. Clearly, there is a need to identify early subcellular changes in the prostate, such as in DNA, associated with aging, primary tumor development, and the progression to metastasis.

We have reported (4) that the ratio of mutagenic 8-hydroxy to putatively nonmutagenic ring-opened purine lesions in prostate DNA increased ≈3-fold in men between the ages of 16 and 83. In addition, a related study (5) showed a substantial accumulation of 8-hydroxyguanine (8-OH-Gua) in the liver, kidney, and intestine as rats aged. It was later reported (6) that levels of this base lesion increased 186% in the livers of rats between 5 and 30 months. Several factors have been suggested to explain age-related increases in DNA base lesion concentrations in various tissues, among them reduced enzyme repair (7–10), a slow loss of DNA nuclease activity (5), and an increase in base oxidation (11). These findings focus attention on the importance of age-related free radical damage to DNA and the putative relationship to cancer. Although a host of factors [e.g., gene mutations (12) and hypermethylation (13)] probably contribute to age-related changes in prostate DNA, there is limited understanding about how many of these events influence the transformation of normal cells to cancer in older men or contribute to metastasis.

We have also reported (14), using statistical models of Fourier transform-IR (FT-IR) spectra, virtually perfect discrimination between the DNA of histologically normal prostates, benign prostatic hyperplasia, and prostatic adenocarcinoma. Significant spectral (structural) differences identified in the bases and backbones of each of the DNA groups contributed to this discrimination. In this regard, our previous study on breast cancer (15) demonstrated that a multivariate model of FT-IR spectral data was able to discriminate between the DNA of nonmetastasizing and metastasizing invasive ductal carcinomas of the female breast.

We have extended our interest in age-related structural changes in DNA as they relate to the development of prostate cancer. We now demonstrate that a relatively high proportion of older men have a DNA structural profile that closely resembles that of primary tumors. Furthermore, structural comparisons of the DNA from primary prostate tumors and metastasizing primary tumors allowed us to develop a statistical model having the potential to determine metastasis by examining the primary tumor.

Materials and Methods

Tissue Acquisition. With Institutional Review Board approval, we obtained frozen (−80°C) prostate tissues from the peripheral zone where about two-thirds of tumors occur (3). Forty-nine prostate samples were provided by the following donors: Baylor College of Medicine Specialized Program of Research Excellence tissue bank project (n = 32); Washington Pathology Consultants, Seattle (n = 10); Northwest Tissue Center, Seattle (n = 6); and the Cooperative Human Tissue Network, Pittsburgh (n = 1). The samples comprised the following groups: (*i*) histologically normal tissues from healthy individuals ranging in age from 16 to 80 years (n = 21); (*ii*) tissues from patients with primary prostate cancer, ages 50–75, comprising microscopically isolated tumor tissues (n = 9); (*iii*) microscopically isolated metastasizing primary tumor tissues (based on confirmed distant metastases), ages 55–71 (n = 11); and (*iv*) microscopically isolated distant metastases, ages 48–73, obtained from lymph nodes (n = 7) and the pelvic region (n = 1). All samples were from different individuals.

Abbreviations: FT-IR, Fourier transform-IR; GC-MS, gas chromatography-MS; 8-OH-Gua, 8-hydroxyguanine; 8-OH-Ade, 8-hydroxyadenine; Fapy, formamidopyrimidine; FapyGua, fapyguanine; PC, principal component.

†To whom correspondence should be addressed. E-mail: dmalins@pnri.org.

MEDICAL SCIENCES

Table 1. *P* values for *t* tests of PC scores from sets of FT-IR spectra

Comparisons	PCs									
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Younger normal vs. older normal	0.7	0.09	0.01	1.0	0.5	<0.001	<0.001	0.1	0.2	0.7
Older normal vs. primary tumor	0.6	0.7	0.6	0.8	0.2	0.04	0.9	0.1	1.0	0.7
Primary tumor vs. metast. primary tumor	0.7	0.3	0.2	0.1	0.003	0.8	<0.001	0.7	0.01	0.1
Primary tumor vs. distant metastases	0.3	0.1	<0.001	0.2	0.4	0.04	0.02	<0.001	0.4	0.01

DNA from the following prostate tissues were compared: histologically normal from younger (ages 16–36; $n = 9$) and older (ages 55–80; $n = 12$) men, primary tumors ($n = 9$), metastasizing primary tumors ($n = 11$), and distant metastases ($n = 8$).

Isolation of High Purity Tumor Tissues. Frozen tumor slices inked with different colors to denote posterior or anterior and right or left aspects of the slices were provided to our pathologist (E.A.B.). Two similarly inked, matching glass microscope slides of hematoxylin/eosin sections from the adjacent slices of the tumor were also provided. Tumor foci were located on the "sandwich" hematoxylin/eosin slides by dotting the tumor periphery, then extrapolated to the frozen slice so that the portion containing relatively pure tumor could be dissected. Each isolated tumor tissue was estimated to be ≈90% pure.

DNA Extraction. DNA (≈50 μg) was extracted from each prostate tissue (70–100 mg) with Qiagen (Chatsworth, CA) 100/G Genomic tips by using a modification of the Qiagen extraction procedure. The DNA was passed through a 5.0-μ Cameo 30N filter (Osmonics, Minnetonka, MN) before precipitation. The DNA was then washed three times with ice-cold 70% ethanol. The Qiagen procedure is an ion-exchange system and does not constitute a source for artifactual oxidation of purines during extraction. In preparation for FT-IR spectral and gas chromatography–MS (GC-MS) analyses, the DNA was dissolved in 10–40 μl (depending on the size of the pellet) of optima grade distilled water (Fisher Scientific).

FT-IR Spectroscopy. A 0.2-μl aliquot of the DNA solution was spotted directly on a BaF_2 plate and allowed to spread, forming an outer ring that contained the DNA. Two separate spots (splits) were created for each DNA sample. The spots were allowed to dry. Spotting was repeated until the ring was at least 100 μ wide, the width of the aperture of the microscope spectrometer (System 2000, Perkin–Elmer). The plate was then placed in a lyophilizer for 1 h to completely dry the DNA. Initially, a background energy reading (percent transmittance) was determined from a blank area of the BaF_2 plate. Energy readings were then taken at various points around the ring, and the points for spectral determination were selected where the energy readings were 15–25% less than the background energy (optimally close to 15% less). Ten spectral determinations were made around each of the two rings per sample and the percent transmittance values were converted (Fourier-transformed) into absorbance values. Each spectrum was baselined (the mean absorbance across 11 wavenumbers, centered at the minimum absorbance value between 2,000 and 1,700 cm^{-1}, was subtracted from the total absorbance at each wavenumber) and then normalized (the entire baselined spectral absorbances were divided by the mean between 1,750 and 760 cm^{-1}). Baselining and spectral normalization adjusted for the optical characteristics of each sample (e.g., related to film thickness). The mean absorbance value was then determined at each integer wavenumber (1,750 to 700 cm^{-1}). The average coefficient of variation for all spectra obtained was 1.9%, with a maximum of 4.8%.

GC-MS. GC-MS was performed as reported (4, 16, 17). In short, DNA (≈20 μg) was hydrolyzed with 150 μl of 60% formic acid in Reacti-vials (Pierce). Hydrolysates were lyophilized and then derivatized by using 50 μl of bis(trimethylsilyl)trifluoroacetic acid containing 1% trimethylchlorosilane and acetonitrile (4:1, vol/vol). In a study of the chemical oxidation of guanine (P.M.J., unpublished results), no evidence was found for a significant increase in 8-OH-Gua by GC-MS under the conditions described. These findings are consistent with those reported recently (18–20). GC-MS analysis of the samples was conducted by using a Hewlett–Packard model 6890 gas chromatograph with a Hewlett–Packard model 5973 mass spectrometer. The purine base lesions quantified were 8-OH-Gua, 8-hydroxyadenine (8-OH-Ade), 2,6-diamino-4-hydroxy-5-formamidopyrimidine (Fapyguanine; FapyGua), and 4,6-diamino-5-formamidopyrimidine (Fapyadenine) (21). The concentrations were reported as base lesions per 10^5 parent bases (4).

Statistical Analyses. The representative FT-IR spectrum of each sample was obtained by determining the mean of 20 spectral determinations (10 per split, as described above). It is often difficult to visually discern subtle differences between mean DNA spectra for different groups of samples. Therefore, a *t* test was performed to determine the statistical significance (*P* value) of differences at each wavenumber between mean absorbance values for each pair of groups (e.g., primary tumor vs. metastasizing primary tumor DNA). Although the *P* values at almost 1,000 consecutive wavenumbers are not statistically independent, regions of the spectrum with $P < 0.05$ are likely to have real structural differences between the groups.

Principal components (PCs) analysis, which involves $\approx 10^6$ correlations between spectral absorbances per wavenumber, integrates different properties of the spectra (e.g., varying peak heights, peak locations, and various combinations thereof). PCs analysis was performed on the mean spectrum of each sample, resulting in 10 PC scores per sample (22, 23). Significant differences in the PC scores between tissue groups were determined by using *t* tests (Table 1). PCs with significant differences ($P < 0.05$) were used to construct two-dimensional scatter plots. The separation of sample clusters in the plots signifies that the groups are structurally dissimilar (15, 22). Significant differences ($P < 0.05$) in PCs of primary prostate tumors and metastasizing primary tumors were used in logistic regression analysis to predict the probability of metastatic cancer, and the sensitivity and specificity of prediction were calculated. The relationship between PCs and age was evaluated by using Spearman correlation coefficients.

The Levene test (24) was used to find significant differences in variance between groups of base lesion concentrations; *t* tests were performed to identify significant differences in the mean concentrations of base lesions between the DNA of tissue groups. The equal variance *t* test was used, unless the Levene test revealed significant differences in variance, in which case the unequal variance version was used. The relationship between base lesion concentrations and age was established by using Spearman correlation coefficients.

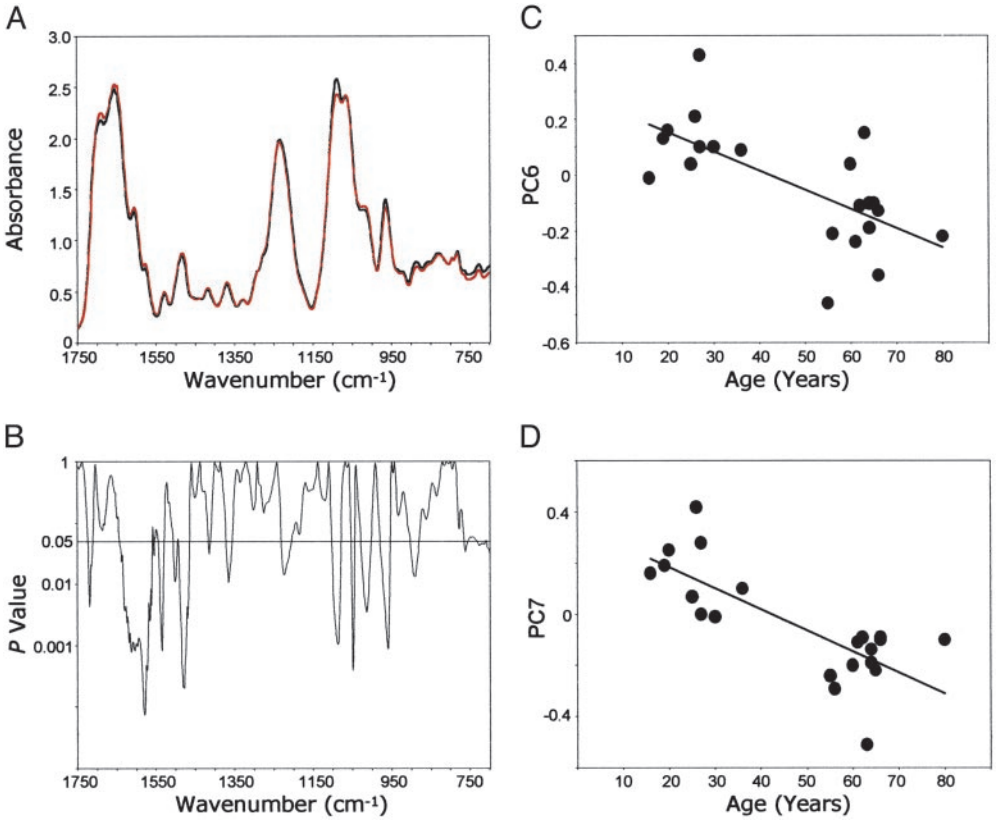

Fig. 1. Age-related differences in FT-IR spectra of DNA from histologically normal prostate tissues from two groups of men, ages 16–36 (mean = 25.1; $n = 9$) and 55–80 (mean = 63.5; $n = 12$). (*A*) Mean DNA spectra for the younger (black) and the older (red) groups. (*B*) Comparison of differences (*P* values) between the mean spectra at each wavenumber. (*C*) Relationship between PC6 and age. (*D*) Relationship between PC7 and age.

Results and Discussion

Age-Related Changes in DNA of the Normal Prostate. Aging is associated with a variety of detrimental changes in organisms at virtually all levels of biological organization (25). Many of these changes lead to degenerative diseases, cancer, and ultimately death. The present findings suggest that substantial age-related changes in the structure of DNA are implicated in the development of prostate cancer in older men.

DNA from histologically normal prostate tissues of men ages 16–36 ($n = 9$) and 55–80 ($n = 12$) was analyzed by FT-IR spectroscopy. A comparison between the mean DNA spectra of the younger and the older men (Fig. 1*A*) revealed regions with significant structural differences between the groups ($P < 0.05$; Fig. 1*B*). These differences, reflecting complex changes in DNA structure, occurred in the bases and backbone and extended over 34% of the spectral range. ($P < 0.05$ would be expected for only ≈5% of the wavenumber range if all of the group differences were random.) For example, there were differences in mean values attributed to $C_4{=}O$ stretching vibrations (assigned to thymine residues) represented by the absorbance at 1,720 cm^{-1} (Fig. 1 *A* and *B*). [The assignments of DNA structure to spectral vibrations are based on the literature (26, 27).] There were also differences in NH_2 bending and C=N stretching vibrations evident at ≈1,606 cm^{-1} (attributed to adenine residues) (Fig. 1 *A* and *B*). In addition, the mean differences in bands between 1,550 and 1,300 cm^{-1} (Fig. 1 *A* and *B*) reflect in-plane base vibrations of the NH and CH deformation modes. Spectral vibrations between ≈1,700 and 1,350 cm^{-1} are attributed to alterations in base structures (e.g., NH_2) and are known to induce changes in conformational properties, such as those associated with vertical base stacking interactions. The strong peak at 1,088 cm^{-1} (Fig. 1*A*) is assigned to symmetric stretching vibrations of the PO_2^- backbone component. The shoulder at 1,021 cm^{-1} and the peak at 960 cm^{-1} are two of several bands in this spectral region assigned to ribose-phosphate main-chain vibrations. These comparisons of mean spectral properties provide persuasive evidence that the base and backbone structures of prostate DNA are progressively altered as a function of increasing age.

To better understand the relationship between age and DNA changes in the histologically normal prostate, PC scores derived from the DNA spectra for each of the men were plotted against their ages (16–80 years; Fig. 1 *C* and *D*). The PCs selected (PC6 and PC7) were based on *t* tests (each *P* value < 0.001) shown in Table 1. The correlation coefficients between age and PC scores were as follows: PC6, −0.622, $P = 0.003$ and PC7, −0.679, $P = 0.001$.

Table 1 also shows many significant group differences in PC scores. There are 12 of 40 *P* values <0.05, whereas only two would be expected by chance if the PC scores were not related to the groups. Also of note, the higher PC scores, such as PC5 and beyond, tend to have more of the significant *P* values than the lower PC scores. In general, the higher the PC score, the more subtle are the features described, and PC scores 5 and higher would tend to represent extremely subtle differences among spectra, explaining why differences in group mean spectra are often not visually apparent. Also, the comparison between samples from older normal men and from primary tumors in Table 1 shows only one PC with a (mildly) significant difference (PC6), whereas all other group comparisons show three to five PC scores with significant differences. All of the groups have a roughly similar size ($n = 8$–11), so the fewer

Table 2. Mean DNA base lesion concentrations with *P* values for differences between histologically normal prostates from men ages 16–27 and men ages 55–66

	Younger normal (n = 8)		Older normal (n = 11)		*P* value	
Base lesion	Mean	Variance	Mean	Variance	Mean*	Variance†
8-OH-Gua/10^5 Gua	233	1,087	310	12,251	0.05	0.02
FapyGua/10^5 Gua	58	271	101	1,584	<0.01	0.02
8-OH-Ade/10^5 Ade	73	226	118	2,305	0.01	0.05
FapyAde/10^5 Ade	40	135	43	165	0.58	0.98

The data for the 36- and 80-year-olds were not used in these groups because the base lesion concentrations were >3 SDs from the mean.
**t* test.
†Levene test.

differences detected between older men and primary cancer are not caused by lower statistical power, but are more likely caused by genuine similarity of the mean spectra of these two groups.

Reactive oxygen species, notably the ·OH, have been implicated in reactions with the base (21, 28) and backbone (29) structures of DNA. Recent evidence (30) has shown that the modification of DNA bases by this radical (i.e., represented by the introduction of a single 8-oxo group into a 25-base strand) induces changes in the conformational structure of the phosphodiester-deoxyribose moiety. In addition, the ·OH is known to react directly with the phosphodiester-deoxyribose structure, such as by abstracting an H atom from the sugar component (31).

In a previous study (4), we found that the proportion of mutagenic 8-OH-purine lesions to putatively nonmutagenic Fapy lesions in prostate DNA increased significantly as men aged. In the present work, significantly higher mean concentrations of 8-OH-Gua (P = 0.05), 8-OH-Ade (P = 0.01), and FapyGua ($P < 0.01$) were found in prostate DNA of the older group of men compared with the younger group (Table 2). These elevated levels of base lesions in the DNA of the older men are assumed to be a factor contributing to the structural changes found between the age groups by FT-IR spectroscopy. Moreover, the variance values for the mean concentrations of 8-OH-Gua, 8-OH-Ade, and FapyGua were substantially higher in the older age group (Table 2). These differences in means and variances between the younger and older men are also evident in the plots of base lesion concentrations with age (Fig. 2). There were no significant differences in either the mean or variance for Fapyadenine (Table 2). The Spearman correlation coefficients between base lesion concentrations and age were 0.52 (P = 0.015) for 8-OH-Gua, 0.55 (P = 0.01) for 8-OH-Ade, and 0.50 (P = 0.02) for FapyGua. In a previous study (32), increased variance in mutagenic base lesions in tumor-prone tissues was hypothesized to be an obligatory precursor to clonal selection and subsequent tumor formation.

Most significant is the fact that the base lesion concentrations for a high proportion of the older men substantially exceeded the highest base lesion values for the younger men (Fig. 2). Using the 8-OH-Gua concentrations as an example, the values ranged from ≈150 to 300 base lesions per 10^5 parent bases for the younger men and ≈150–500 for the older men. The slope of the line shows the increase in mean base lesion concentrations for men aged 60, compared with men aged 20, was 60% for 8-OH-Gua, 80% for 8-OH-Ade, and 160% for FapyGua. Some men, however, showed no increase in base lesion concentrations with age (Fig. 2). The overall implication is that some older men maintain the base lesion concentrations characteristic of the younger men, whereas others develop substantially elevated concentrations, notably of the mutagenic purine lesions. The increases in the mutagenic purine lesions in the older men may well increase their risk for prostate cancer (4).

Similarities Between the DNA of Older Men and the DNA of Primary Prostate Tumors. FT-IR spectroscopy showed that the DNA from a substantial subset of the older men was structurally very similar to the DNA of primary prostate tumors. When the mean DNA spectrum of the older men was compared with the mean spectrum of primary prostate tumors, no significant differences were evident over the entire spectral range. However, despite the inability to discriminate between these mean spectra, logistic regression analysis based on PC6 (P = 0.04; Table 1) revealed a marked discrimination between the two groups (Fig. 3). Five of 12 of the points representing the older men fell within the group for the primary tumor DNA; that is, 42% of the older men had a DNA phenotype indistinguishable from that of primary prostate cancer. This finding suggests that these older men with the cancer-like phenotype had a relatively high cancer risk. By contrast, 58% of the points for the older men fell below the tumor group and were completely separated from it. It is tempting to speculate that these men had a relatively low risk for

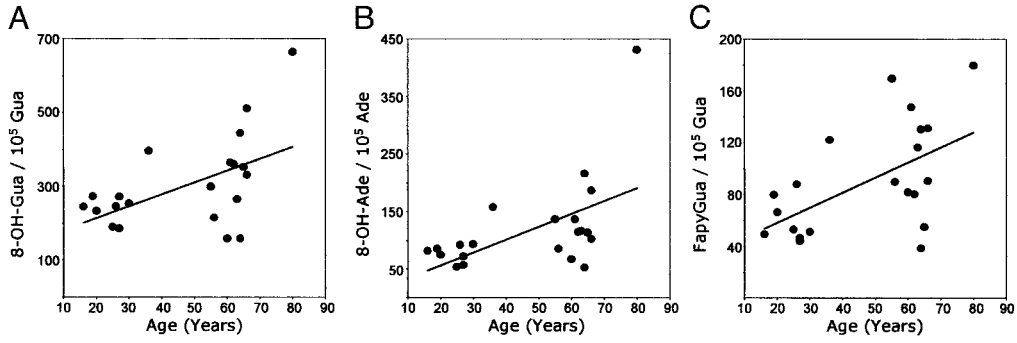

Fig. 2. Concentrations of 8-OH-Gua (*A*), 8-OH-Ade (*B*), and FapyGua (*C*) in DNA of histologically normal prostate tissues from men ages 16–80 (n = 21) vs. age. See text for details.

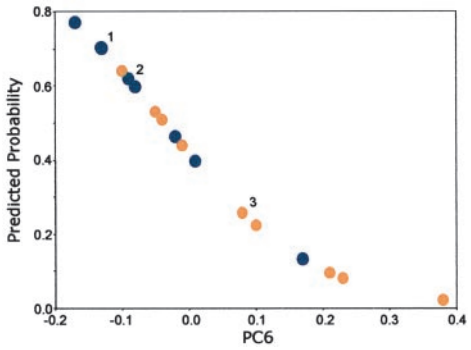

Fig. 3. Predicted probability of cancer based on a logistic regression model. The model uses PC6 from FT-IR spectra of DNA from histologically normal prostates of men (ages 55–80; $n = 12$; orange) and DNA from primary prostate tumors of men (ages 50–75; $n = 9$; blue). Overlapping points: 1, two primary tumors and one normal; 2, two primary tumors; and 3, three normals.

prostate cancer, which is consistent with the epidemiological evidence showing that some older men never develop this disease (33). [The high variance values for base lesion concentrations (Table 2) and the small number of samples for the putatively high and low cancer risk sub groups (Fig. 3) provided little statistical power for detecting differences in base lesion concentrations between these groups.]

We have shown that FT-IR microscope spectroscopy is capable of identifying subtle age-related changes in DNA that, in some men, mimic the prostate cancer phenotype. Thus, the results obtained

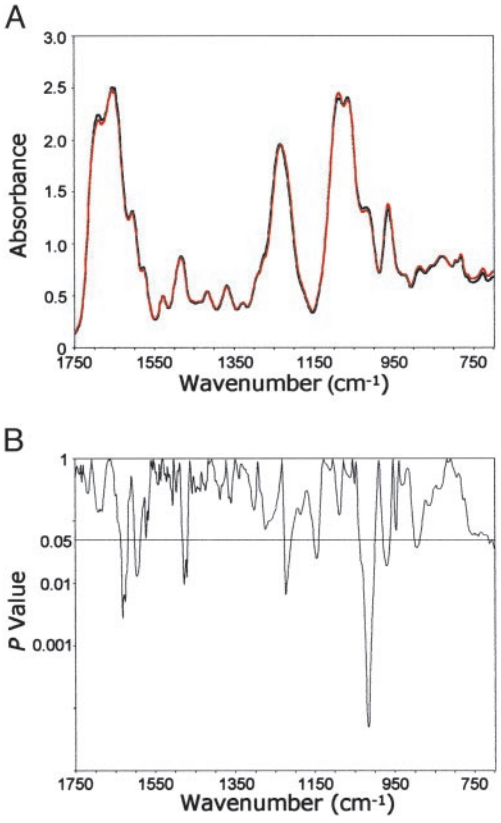

Fig. 4. Comparison of FT-IR spectra of DNA from primary prostate tumor tissues ($n = 9$) and metastasizing primary prostate tumor tissues ($n = 11$). (*A*) Mean DNA spectra for the primary prostate tumor (black) and metastasizing primary tumor (red) tissues. (*B*) Comparison of differences (*P* values) between the mean spectra at each wavenumber.

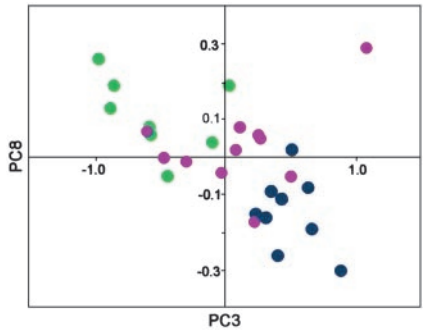

Fig. 5. Plot of PC3 vs. PC8 from FT-IR spectra of DNA from primary prostate tumor tissues ($n = 9$; blue), metastasizing primary prostate tumor tissues ($n = 11$; pink), and distant metastases of prostate tumors ($n = 8$; green).

provide a promising basis for the development of cancer risk models. A prime example of a potentially high risk factor may well be a close structural similarity between a man's DNA and that established for primary prostate cancer. However, future models based on larger sample numbers and genetic and lifestyle factors may additionally allow for men at risk to be identified before they develop the characteristic cancer phenotype. Because free radicals are apparently a significant component in inducing the structural alterations in prostate DNA (4), intervention with antioxidants (e.g., in the diet) may slow this process and thus extend the age at which the cancer phenotype is reached. A basis for this conclusion has been suggested from dietary intervention studies with natural products, such as those containing the antioxidants and radical-trapping agents lycopene (34, 35), tocopherols (36), and polyphenols (37, 38) that reportedly inhibit prostate cancer development.

Predicting Metastasis of a Primary Tumor. Because we were able to discriminate between the DNA of younger and older men, we were interested in determining whether it would be possible to discriminate between the DNA of a primary tumor and one that had metastasized. The ability to determine whether a primary prostate tumor has progressed to the metastatic state has considerable clinical potential and is of prime importance in avoiding a fatal outcome (3). However, we are unaware of any method for determining whether a primary prostate tumor is metastasizing until distant metastases have been observed (39), thus considerably reducing prospects for successful intervention (3). Therefore, a distinct advantage in patient treatment would be the development of a practical and reliable means for determining whether a prostate

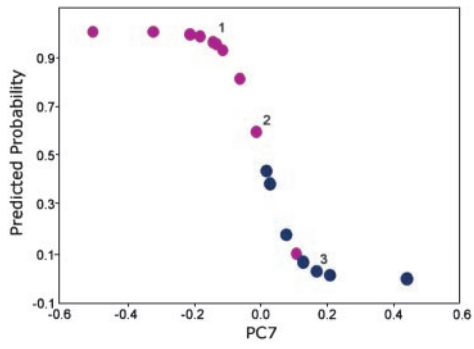

Fig. 6. Predicted probability of metastatic cancer based on a logistic regression model. The model uses PC7 from FT-IR spectra of DNA from primary prostate tumors ($n = 9$; blue) and metastasizing primary prostate tumors ($n = 11$; pink). Overlapping points: 1, three metastasizing primary tumors; 2, one primary tumor and one metastasizing primary tumor; and 3, two primary tumors.

tumor is metastasizing before any distant metastases have been discovered. We propose that the use of FT-IR spectroscopy as described offers an attractive means for identifying those patients in need of aggressive therapy.

As anticipated, FT-IR spectroscopy effectively distinguished between the DNA of primary tumors and the DNA of distant metastases. In fact, the significant differences between the two spectral means spanned over 42% of the entire spectral range and involved a number of structural components comprising the bases and backbone. However, this capability was not pursued extensively given that these two tumor states are readily distinguished histologically. Rather, we turned our attention to the more important issue of distinguishing between the DNA of primary tumors and the DNA of metastasizing primary tumors.

We found significant structural differences over 15.2% of the spectral range when the mean spectra of DNA from each of these tissue types were compared (Fig. 4 *A* and *B*). These differences occurred in the bases and vertical base-stacking interactions, as reflected in the significant *P* values between ≈1,650 and 1,480 cm^{-1} and in the region assigned to vibrations of the phosphodiester-deoxyribose moiety (≈1,227 to 900 cm^{-1}). Examples of these backbone differences include antisymmetric PO_2^- stretching vibrations (≈1,225 cm^{-1}) and a shoulder at ≈1,020 cm^{-1} reflecting ribose-phosphate main-chain vibrations (26, 27).

Considering that the primary tumor, metastasizing primary tumor, and the distant metastases are integrally related components of the metastatic process, we determined whether the DNA from these three tumor stages could be separated using PCs analysis. PC scores were obtained for all three DNA groups combined. *P* values were determined by using *t* tests between PC scores for the DNA of the primary tumors and the distant metastases. Strikingly, a plot of the two most significant PCs, PC3 and PC8 ($P < 0.001$; Table 1), revealed an almost perfect discrimination between the three DNA groups (Fig. 5). Of particular interest was the fact that the cluster of points for the DNA of the metastasizing primary tumors occurred as an intermediate between the primary tumor DNA and that of the distant metastases. The Spearman correlation coefficient between PC3 and PC8 for all three groups combined was -0.48 with $P = 0.01$.

The impressive discrimination obtained (Fig. 5) provided a strong justification for conducting a logistic regression analysis for the primary prostate tumor DNA and the metastasizing primary tumor DNA. The model selected was based on PC7 ($P < 0.001$; Table 1). At 50% predicted probability for metastasis, we found that the sensitivity was 91% and the specificity was 89% (Fig. 6). The ability to discriminate with such accuracy the DNA of the primary tumors and the DNA of the metastasizing primary tumors is highly promising for predicting whether a primary tumor has metastasized. In a clinical application, a fine needle biopsy yielding a minimum of 1 μg of DNA would be sufficient for this determination. We anticipate that future studies, using a larger number of samples, would provide added strength to support the present findings.

Conclusions

Evidence has been presented to show that FT-IR microscope spectroscopy reveals subtle structural changes in DNA associated with both aging and cancer development in the prostate. In a relatively high proportion of the normal men studied, age-related changes in nonmalignant tissues were associated with the development of a DNA phenotype with features very similar to that of primary prostate tumors. We hypothesize that this cancer-like phenotype could be used to identify men at high risk for prostate cancer. We further hypothesize that models based on age-related structural changes preceding the manifestation of this phenotype may also serve to predict cancer risk. Of additional importance, we have demonstrated the ability of microspectroscopy to identify changes in DNA that occurred after primary prostate tumors had displayed the characteristic features of metastasis. This finding constitutes a novel approach for the early detection of metastasis using discrimination models, such as logistic regression, as described.

We thank Drs. Ercole L. Cavalieri, Sandra J. Gunselman, Karl Erik Hellström, Elaine A. Ostrander, Eleanor G. Rogan, Edward L. Weber, and Donald Withers for their helpful comments and Dr. Virginia M. Green for editorial assistance. Naomi Gilman and Brenda Duer provided technical assistance. We also thank the Cooperative Human Tissue Network, the Northwest Tissue Center, Washington Pathology Consultants, and the Specialized Program of Research Excellence at Baylor College of Medicine (funded by National Cancer Institute Grant CA58204) for providing tissues and pathology data, with special thanks to Mohammed Sayeeduddin. This study was supported by National Cancer Institute Grant CA79690.

1. Stanford, J. L., Stephenson, R. A., Coyle, L. M., Cerhan, J., Correa, R., Eley, J. W., Gilliland, F., Hankey, B., Iolonel, L. N., Kosary, C., *et al*. (1999) *Prostate Cancer Trends 1973–1995* (National Institutes of Health, Bethesda), Surveillance, Epidemiology, and End Results Program, National Cancer Institute, National Institutes of Health Publ. No. 99–4543.
2. American Cancer Society (2003) *Cancer Facts and Figures* (American Cancer Society, Atlanta).
3. Kirby, R. S., Christmas, T. J. & Brawer, M. K. (1996) *Prostate Cancer* (Mosby, London).
4. Malins, D. C., Johnson, P. M., Wheeler, T. M., Barker, E. A., Polissar, N. L. & Vinson, M. A. (2001) *Cancer Res.* **61,** 6025–6028.
5. Fraga, C. G., Shigenaga, M. K., Park, J. W., Degan, P. & Ames, B. N. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4533–4537.
6. Schmerold, I. & Niedermuller, H. (2001) *Exp. Gerontol.* **36,** 1375–1386.
7. DePinho, R. A. (2000) *Nature* **408,** 248–254.
8. Ames, B. N., Shigengaga, M. K. & Gold, L. S. (1993) *Environ. Health Perspect.* **101,** Suppl. 5, 35–44.
9. Reid, T. M. & Loeb, L. A. (1993) *Mutat. Res.* **289,** 181–186.
10. Wei, Q., Matanoski, G. M., Farmer, E. R., Hedayati, M. A. & Grossman, L. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 1614–1618.
11. Osterod, M., Hollenbach, S., Hengstler, J. G., Barnes, D. E., Lindahl, T. & Epe, B. (2001) *Carcinogenesis* **22,** 1459–1463.
12. Loeb, L. A., Loeb, K. R. & Anderson, J. P. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 776–781.
13. Oakes, C. C., Smiraglia, D. J., Plass, C., Trasler, J. M. & Robaire, B. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 1775–1780.
14. Malins, D. C., Polissar, N. L. & Gunselman, S. J. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 259–264.
15. Garcia-Closas, M., Hankinson, S. E., Ho, S., Malins, D. C., Polissar, N. L., Schaefer, S. N., Su, Y. & Vinson, M. A. (2000) *J. Natl. Cancer Inst. Monogr.* **27,** 147–156.
16. Rodriguez, H., Jurado, J., Laval, J. & Dizdaroglu, M. (2000) *Nucleic Acids Res.* **28,** e75: i–vii.
17. Dizdaroglu, M. (1994) *Methods Enzymol.* **234,** 3–16.
18. Sentürker, S. & Dizdaroglu, M. (1999) *Free Radical Biol. Med.* **27,** 370–380.
19. Dizdaroglu, M. (1998) *Free Radical Res.* **29,** 551–563.
20. England, T. G., Jenner, A., Aruoma, O. I. & Halliwell, B. (1998) *Free Radical Res.* **29,** 321–330.
21. Malins, D. C., Holmes, E. H., Polissar, N. L. & Gunselman, S. J. (1993) *Cancer* **71,** 3036–3043.
22. Malins, D. C., Polissar, N. L., Su, Y., Gardner, H. S. & Gunselman, S. J. (1997) *Nat. Med.* **3,** 927–930.
23. Malins, D. C., Polissar, N. L., Nishikida, K., Holmes, E. H., Gardner, H. S. & Gunselman, S. J. (1995) *Cancer* **75,** 503–517.
24. Zar, J. H. (1999) *Biostatistical Analysis* (Prentice–Hall, Upper Saddle River, NJ).
25. Troen, B. R. (2003) *Mt. Sinai J. Med.* **70,** 3–22.
26. Shimanouchi, T. & Tsuboi, M. (1964) in *Advances in Chemical Physics: The Structure and Properties of Biomolecules and Biological Systems*, ed. Duchesne, J. (Wiley, London), Vol. 7, pp. 436–498.
27. Tsuboi, M. (1969) *Appl. Spectrosc. Rev.* **3,** 45–90.
28. Dizdaroglu, M. (1992) *Mutat. Res.* **275,** 331–342.
29. Dizdaroglu, M., Jaruga, P., Birincioglu, M. & Rodriguez, H. (2002) *Free Radical Biol. Med.* **32,** 1102–1115.
30. Malins, D. C., Polissar, N. L., Ostrander, G. K. & Vinson, M. A. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 12442–12445.
31. von Sonntag, C. (1987) *The Chemical Basis of Radiation Biology* (Taylor and Francis, New York).
32. Malins, D. C., Polissar, N. L., Schaefer, S., Su, Y. & Vinson, M. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 7637–7642.
33. Bonafè, M., Barbi, C., Storci, G., Salvioli, S., Capri, M., Olivieri, F., Valensin, S., Monti, D., Gonos, E. S., De Benedictis, G. & Franceschi, C. (2002) *Exp. Gerontol.* **37,** 1263–1271.
34. Bowen, P., Chen, L., Stacewicz-Sapuntzakis, M., Duncan, C., Sharifi, R., Ghosh, L., Kim, H. S., Christov-Tzelkov, K. & van Breemen, R. (2002) *Exp. Biol. Med. (Maywood)* **227,** 886–893.
35. Chen, L. Stacewicz-Sapuntzakis, M., Duncan, C., Sharifi, R., Ghosh, L., van Breemen, R., Ashton, D. & Bowen, P. E. (2001) *J. Natl. Cancer Inst.* **93,** 1872–1879.
36. Giovannucci, E. (2000) *J. Natl. Cancer Inst.* **92,** 1966–1967.
37. Gupta, S., Hastak, K., Ahmad, N., Lewin, J. S. & Mukhtar, H. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 10350–10355.
38. Gupta, S., Ahmad, N., Mohan, R. R., Husain, M. M. & Mukhtar, H. (1999) *Cancer Res.* **59,** 2115–2120.
39. National Cancer Institute (2000) *Cancer Facts* (National Institutes of Health, Bethesda), Vol. 2003.