Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 259–264, January 1997
Medical Sciences

Models of DNA structure achieve almost perfect discrimination between normal prostate, benign prostatic hyperplasia (BPH), and adenocarcinoma and have a high potential for predicting BPH and prostate cancer

(radical-induced DNA damage/cancer prediction and screening/infrared spectroscopy)

DONALD C. MALINS*, NAYAK L. POLISSAR†‡, AND SANDRA J. GUNSELMAN*

*Molecular Epidemiology Program, Pacific Northwest Research Foundation, 720 Broadway, Seattle, WA 98122; †The Mountain-Whisper–Light Statistical Consulting, Seattle, WA 98112; and ‡Department of Biostatistics, University of Washington, Seattle, WA 98195

Contributed by Donald C. Malins, November 7, 1996

ABSTRACT **In our previous studies of DNA, wavenumber–absorbance relationships of infrared spectra analyzed by principal components analysis (PCA) were expressed as points in space. Each point represented a highly discriminating measure of structural modifications that altered vibrational and rotational motion, thus changing the spatial orientation of the points. PCA/Fourier transform-infrared technology has now provided a virtually perfect separation of clusters of points representing DNA from normal prostate tissue, BPH, and adenocarcinoma. The findings suggest that the progression of normal prostate tissue to BPH and to prostate cancer involves structural alterations in DNA that are distinctly different. The hydroxyl radical is likely a major contributor to these structural alterations, which is consistent with previous studies of breast cancer. Models based on logistic regression of infrared spectral data were used to calculate the probability of a tissue being BPH or adenocarcinoma. The models had a sensitivity and specificity of 100% for classifying normal vs. cancer and normal vs. BPH, and close to 100% for BPH vs. cancer. Thus, the PCA/Fourier transform-infrared technology was shown to be a powerful means for discriminating between normal prostate tissue, BPH and prostate cancer and has considerable promise for risk prediction and clinical application.**

Prostate cancer is a leading cause of death (1). Thus, there is a keen interest in the etiology of this disease, as well as in the development of techniques for predicting its occurrence at early stages of oncogenesis. Little is known about the etiology of prostate cancer, the most prevalent form being adenocarcinoma (1). However, several studies have focused on inactivation of the tumor suppressor gene TP53 and altered DNA methylation patterns (2) as possible factors. In addition, free radicals, arising from redox cycling of hormones, have recently been implicated in prostate cancer (3). This is consistent with evidence showing that the hydroxyl radical (·OH) produces mutagenic alterations in DNA, such as 8-hydroxyguanine (8-OH-Gua) and 8-hydroxyadenine (8-OH-Ade), that have been linked to carcinogenesis in a variety of studies (4–14). Moreover, significant base misincorporation in DNA synthesis is known to be associated with the formation of these 8-OH-derivatives (5, 6, 13). The ·OH is believed to arise from H_2O_2 via the Fe^{2+}-catalyzed Fenton reaction. H_2O_2 readily crosses the nuclear membrane and enters the nucleus where it is believed to be converted to the ·OH (15), which can damage DNA (16).

In our studies of breast cancer, 2-fold higher concentrations of 8-OH-Ade and 8-OH-Gua were found in primary invasive ductal carcinoma DNA compared with normal breast DNA (9). Similarly, a 2-fold increase in 8-OH-Ade was found in metastatic invasive ductal carcinoma compared with the primary tumor (11). Despite these findings, virtually no understanding exists of the possible relationship between the ·OH-modification of DNA and prostate cancer.

Prostate tissue may contain areas of BPH, which is not regarded as a premalignant lesion (1), although it often accompanies prostate cancer. The etiology of BPH is unknown, as is its relationship to prostate cancer (1). However, recent evidence with dogs (3) suggests that redox cycling of hormones may be a mechanism for the radical induction of BPH. Moreover, in humans there is evidence showing that increased levels of 8-OH-Gua, 8-OH-Ade, and other ·OH-modified bases are present in BPH tissue compared with adjacent normal tissue (17). This finding is consistent with results with nonneoplastic tissue from a phylogenetically different source (i.e., fish). Positive correlations were obtained between concentrations of 8-OH-Gua and 8-OH-Ade and the incidence of a number of hepatic lesions (e.g., putatively preneoplastic basophilic foci) in cancer-free tissues of English sole from populations prone to liver cancer (18). Thus, the evidence points to a positive association between radical-induced mutagenic base modifications and the formation of nonneoplastic lesions in cancer-prone tissues.

We used a different set of models, based on Fourier transform-infrared (FT-IR) spectroscopy, to show structural alterations in DNA in relation to carcinogenesis (10). The DNA alterations were linked to tumor formation in the human female breast, and the probability of cancer was assessed using logistic regression of statistics which describe the FT-IR wavenumber–absorbance relationships between individual specimens, particularly those relationships associated with the base and phosphodiester–deoxyribose structure (10). Recently, FT-IR spectral models were developed that employ principal components analysis (PCA) of FT-IR spectra (11, 12). This powerful technology allowed spectra to be expressed as points in space, each point being a highly discriminating measure of DNA structure. Changes in the cellular environment (e.g., increased free radicals) can potentially lead to alterations in the vibrational and rotational motion of functional groups of DNA, thus changing the spatial location of the points. We showed in the breast cancer studies (11, 12) that clusters of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

0027-8424/97/94259-6$2.00/0
PNAS is available online at **http://www.pnas.org**.

Abbreviations: PC, principal components; PCA, principal components analysis; FT-IR, Fourier transform-infrared; BPH, benign prostatic hyperplasia; 8-OH-Ade, 8-hydroxyadenine; 8-OH-Gua, 8-hydroxyguanine.

points had different spatial locations, cluster sizes, or both, depending on whether the DNA was derived from normal, primary tumor, or metastatic tumor tissue (12).

We have now applied the PCA/FT-IR technology to DNA derived from the normal prostate, BPH, and adenocarcinoma. Clusters of points representing DNA from each of these tissues were almost completely separated in two-dimensional plots of principal components (PC) scores. This indicates that significant and specific structural modifications in DNA occur in the progression of normal tissue to BPH and normal tissue to prostate cancer, and that the modifications are unique for each of the two progressions. The structural alterations, reflected primarily in spectral regions representing vibrations of the nucleic acids, phosphodiester, and deoxyribose structures, were consistent with previously observed structural changes of DNA in breast cancer (11, 12). The separation and classification of the normal prostate vs. BPH or adenocarcinoma was investigated using logistic regression models of infrared spectra like those applied to breast cancer (11, 12). Similarly, logistic regression models of DNA spectra were used to evaluate the relationship between BPH and prostate cancer.

MATERIALS AND METHODS

Tissue acquisition, DNA isolation, and FT-IR spectroscopy were carried out as described (10). All tissues were kept at −80°C before use, and DNA was maintained under an atmosphere of pure nitrogen during the extraction procedure to avoid oxidation. A total of 31 tissue samples were used. Five samples of prostate tissue obtained from individuals who died by accidents were examined histologically and found to be normal. These served as controls. Eighteen samples of BPH and 8 samples of adenocarcinoma served as test samples, each comprising a portion of the histologically identified lesion. All samples were obtained from the Cooperative Human Tissue Network (Cleveland) together with related pathology data.

For PCA/FT-IR spectral analysis, each spectrum was normalized across the range of 1750 to 700 cm^{-1}, as described (10). This yielded a relative absorbance value for each wavenumber, with a mean of 1.0. Euclidean distance was used to define the difference between a pair of spectra either for the entire spectrum or for a subregion (10, 11). This standard distance measure is defined as the square root of the sum of squared absorbance differences between spectra at each of the wavenumbers considered (e.g., 1051 for the entire spectral region 1750–700 cm^{-1}). The Euclidean distance can also be expressed in a more descriptive form as a percent. The numerator of the percent is the Euclidean distance divided by the square root of the number of wavenumbers for a region. The denominator used here for the percent for any region is the mean normalized absorbance between 1750 and 700 cm^{-1}, which is 1.0 for every case.

PC analysis was used to identify a few variables (components) that capture most of the information in the original, long list of variables (the spectral absorbances at each wavenumber). This reduction in the number of variables is analogous to the process in educational testing whereby many individual test scores, such as in reading and arithmetic, are combined into a single academic performance score. Four PC scores (i.e., four dimensions) were found to be sufficient to describe the 1051 dimensions of the normalized spectra. PC scores were calculated with the grand mean of all spectra subtracted from each spectrum. The nonparametric Spearman correlation coefficient was used to asses the association of PC scores with patient ages and Gleason scores. The nonparametric analysis was used because some of the distributions are skewed or are not normal ("bell-shaped"), which can lead to a bias in statistical significance when estimated from the Pearson correlation coefficient.

Two cases, which were outliers, were omitted from these analyses, leaving 29 cases. The omitted BPH sample and the omitted cancer sample had spectra very different from the included cases. Their Euclidean distances from the most similar spectra were 52% and 41%, respectively. All other spectra differed from their "nearest neighbor" spectrum by at most 21%, with a majority of spectra differing by less than 11%. The two outlier spectra show drastically reduced absorbance in the region around 1650 cm^{-1} representing vibrations of the nucleic acids.

The Kruskal–Wallis and Mann–Whitney tests were used to determine if the three groups had similar diversity, defined as the mean distance of a spectrum to its group centroid. A permutation test was used to determine whether the three groups tended to cluster separately (representing an internal similarity of spectral properties in a group). The distance of each spectrum to its nearest neighbor in its own group (either normal, BPH, or cancer) was calculated, and the mean of these nearest neighbor distances for of all of the spectra was the test statistic. The test was carried out by randomly permuting group membership labels 1000 times and recalculating the test statistic each time. A smaller observed distance to the nearest neighbor than that obtained by random relabeling of groups is an indication of clustering. A nonparametric, rank-based version of this test was carried out by expressing each distance as a rank. For each spectrum, the distances to other spectra were ranked and the permutation test was carried out as described above, but with distances replaced by ranks. The test statistic was a mean rank. Again, a smaller observed mean rank than the mean obtained from random permutation is an indication of clustering. Both the test using distance and the test using ranks were carried out for the entire spectrum, 1750–700 cm^{-1}, and for several subregions.

Finally, logistic regression analysis was used as a model to determine if PC scores could be used to discriminate between pairs of DNA groups (normal vs. BPH, normal vs. cancer, and BPH vs. cancer). The logistic regression analysis yields a risk score, which is a linear combination of PC scores, and a predicted probability of a sample being in one of the two groups considered (e.g., the probability of being BPH when BPH is compared with normal). These predicted probabilities, along with a chosen probability cut point, can be used to classify samples and provide estimates of sensitivity and specificity, or percent of samples correctly classified. For each analysis a cut point was chosen that jointly maximized sensitivity and specificity.

RESULTS AND DISCUSSION

Clustering in PC Plots. PCA/FT-IR spectral analysis yielded four components (four PC scores per case) which explained a total of 90% of the spectral variation over 1051 wavenumbers. That is, most of the features of the 29 spectra could be described by four PC scores (labeled PC1, PC2, PC3, and PC4). The first two PC scores explained 76% of the variation and were adequate for two-dimensional representation (Fig. 1). Fig. 1 shows that the three groups were distinctly clustered. The two outliers omitted from the analysis are also represented on this plot and appear to the right of the main clusters.

The actual distance of the outlier points to other points is larger than that shown in this two-dimensional plot due to differences represented by other dimensions. The permutation test for clustering of groups (1750–700 cm^{-1}) yielded $P = 0.1$, based on the distance measure, and $P = 0.01$ using the nonparametric ranking technique (Table 1). The greater significance obtained by the ranking method arises from the relative isolation of one or two cases from the core of their group (Fig. 1), a configuration which influences the distance measure more than the ranking measure. Using these tech-

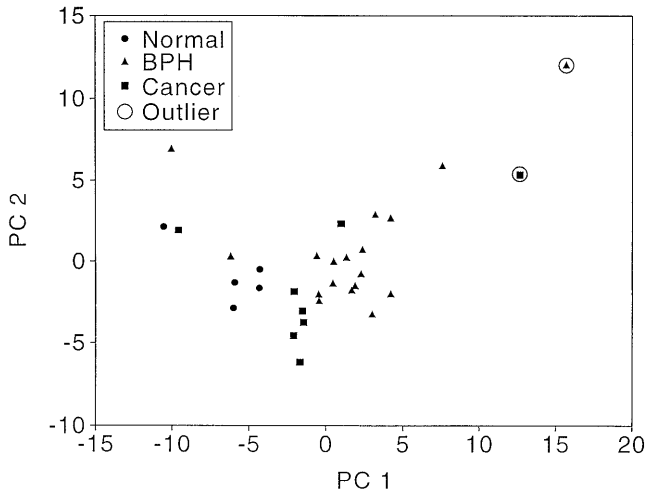

FIG. 1. Two-dimensional PC plot derived by PCA/FT-IR spectral analysis showing distinct clustering of normal, BPH, and prostate cancer points. Notably, both of the groups of prostate lesions occur to the right of the points for the DNA of normal prostate. (See text for details.)

niques, significant clustering was obtained for two regions of the spectrum: 1174–1000 cm^{-1} (assigned to strong stretching vibrations of the PO_2^- and C-O groups of the phosphodiester-deoxyribose structure) and 1499–1310 cm^{-1} (assigned to weak NH vibrations and CH in-plane deformations of the nucleic acids) (19–21). The *P* values for mean distance and mean rank for these regions ranged from 0.02 to <0.001 (Table 1). The significance levels obtained strongly reject the null hypothesis that the observed clustering of the three groups occurred by chance. Overall, the findings indicate that DNA is altered in ways that produce clustering and, consequently, discrimination between normal prostate, BPH, and prostate cancer DNA (Fig. 1 and Tables 1 and 2).

Detailed comparisons were made between the spectra of pairs of groups: normal vs. cancer, normal vs. BPH, and BPH vs. cancer. The statistical significance of differences in mean normalized absorbance between groups was assessed for each wavenumber between 1750 and 700 cm^{-1}, using the unequal variance *t* test (Fig. 2). The plot shows the comparison of the mean spectrum for each of the two groups, as well as the *P* value from the *t* test. The regions with $P \leq 0.05$ represent differences between groups (e.g., normal vs. cancer) which are much less likely to be due to chance than regions with $P > 0.05$. Each of the spectral comparisons between groups shows statistically significant differences in areas of the spectrum assigned to vibrations of the phosphodiester-deoxyribose structure and the nucleic acids. The spectral regions with significant differences in absorbance for the phosphodiester-deoxyribose structure are similar (≈1050–1000 cm^{-1}); however, absorbances associated with the nucleic acids vary among the groups. That is, for the normal vs. cancer comparison, the region of significant difference is primarily ≈1475–1400 cm^{-1}, whereas for the normal vs. BPH comparison it is ≈1600–1500 cm^{-1} (C-O stretching and NH bending vibrations). The comparison for BPH vs. cancer is focused at ≈1500 cm^{-1}. For the normal vs. BPH and BPH vs. cancer comparisons, significant differences are shown between ≈1175 and 1120 cm^{-1}, a region that likely includes symmetric stretching vibrations of the PO_2 group (19–21). The difference in means at all of these spectral regions is apparent from the plots of mean spectra per group in Fig. 2. The structural modifications are pivotal in the spatial distribution of points in the PC plot (Fig. 1) and in the pronounced discrimination between clusters (Table 1). The ·OH is known to produce mutagenic base lesions, such as 8-OH-Gua and 8-OH-Ade (4–14), and also cause damage to deoxyribose by abstracting hydrogens from one or more positions associated with the furanose ring (22). These events can ultimately lead to broadly based genetic instability and strand breaks (23). Recent evidence also indicates that decreased antioxidant levels and increased base modifications occur in BPH tissue compared with adjacent normal prostate (17). Moreover, FT-IR spectral analysis of calf thymus and normal breast DNA exposed for various times to ·OH-generating systems (Fe^{2+}/H_20_2) revealed substantial alterations in areas of the spectrum assigned to vibrations of the nucleic acids and the phosphodiester-deoxyribose moiety (D.C.M., S.J.G., and J. Cramer, unpublished results). Collectively, this evidence supports the proposition that the ·OH is intimately involved in altering the structure of DNA, thus contributing to clustering and discrimination between clusters; however, it is recognized that other factors (e.g., hypermethylation) (2) may also contribute to these alterations. The prostate findings closely resemble those obtained with the female breast (10–12) in which the cancer-related ·OH-modification of DNA was termed radical-induced DNA disorder (RIDD) (11). RIDD also appears to be significant in the etiology of BPH and prostate cancer and constitutes a formidable barrier to overcome in cancer prevention and treatment.

Cluster Diversity. The diversity of the three groups, expressed as the mean distance to the group centroid, did not differ significantly ($P = 0.8$). However, the normal prostate group was slightly less diverse (mean distance = 11.7%) than was the BPH group (mean distance = 14.5%) or prostate cancer group (mean distance = 13.9%). In comparable studies of the female breast, the diversity was not different between normal breast and primary breast cancer DNA. However, significant differences did exist between normal breast and metastatic cancer and primary cancer and metastatic cancer (11, 12). The findings with the normal prostate and primary prostate cancer are consistent with those obtained for the breast, although additional studies are necessary to determine

Table 1. Mean distance to nearest neighbor of same group and permutation test for nonrandom clustering

Spectral region, cm^{-1}	Mean distance*			Mean rank†		
	Observed	Random permutation	*P* value	Observed	Random permutation	*P* value
1750–700	12.2	12.8	0.1	2.0	3.0	0.01
1750–1500	12.3	12.3	0.5	2.4	3.0	0.09
1499–1310	5.9	6.5	0.02	1.6	3.0	<0.001
1309–1175	6.7	6.5	0.7	3.0	3.0	0.5
1174–1000	13.2	15.0	0.02	2.0	3.0	0.01
999–700	6.9	7.4	0.1	2.3	3.0	0.05

Distance is expressed as a percent difference between spectra; 1000 permutations were performed for each spectral subregion.

*Mean Euclidean distance to nearest neighbor in the same group expressed as a percent.

†Mean rank of Euclidean distance of each spectrum to nearest neighbor in the same group.

Table 2. Logistic regression models for probability of BPH (vs. normal), cancer (vs. normal), and cancer (vs. BPH).

Model	Coefficients ± SE					Correct classification rate		
	Intercept	PC1	PC2	PC3	PC4	By group, %	Overall, %	*P* value*
Normal vs. BPH	24.9 ± 0.1	5.2 ± 0.2	5.8 ± 0.04	3.9 ± 0.03		Normal, 100; BPH, 100	100	<0.001
Normal vs. cancer	34.3 ± 0.1	12.0 ± 0.04			−21.0 ± 0.1	Normal, 100; cancer, 100	100	<0.001
BPH vs. cancer	−14.5 ± 8.1	−4.5 ± 2.6	−3.7 ± 2.0		−11.1 ± 6.3	BPH, 88; cancer, 100	92	<0.001

Normal, $n = 5$; BPH, $n = 17$; cancer, $n = 7$. *P* values are based on the null hypothesis that each model is not predictive of group membership. *P* values are calculated from a χ^2 test on change in deviance.
**P* value for the null hypothesis that the probability of a case falling into a specified group is unrelated to the PC scores.

whether DNA from metastatic prostate cancer will show the increased diversity found with metastatic breast cancer. Increased structural diversity generated in primary tumors is likely an important factor in selecting DNA forms that potentially give rise to malignant cell populations, as previously suggested (11).

Group Classification. PC scores can be readily used to classify patients into groups when pairs of groups are compared using logistic regression. The logistic regression model (Table 2) is an equation which yields a risk score, *R*, when the values of the PC scores are inserted into the equation. *R* is transformed to a probability by the following statistical equation: Probability = $\exp(R)/[1 + \exp(R)]$. A cut point is chosen, and if the probability exceeds this cut point, the case would be classified as BPH. The actual cut points are noted below. As shown in Table 2, the model for normal vs. cancer and normal vs. BPH correctly classifies each group 100% and 100% overall (*P* values in each case were <0.001). The correct classification rate for cancer vs. BPH was close to 90% based on a designation of "cancer" for a predicted probability of ≥0.1. (Probability cut-points of 0.15 to 0.41 achieve the same correct classification rates in the BPH vs. cancer comparison). The predicted probabilities based on the models in Table 2 are given in Fig. 3. The individual risk score is based on the appropriate PC model (Table 2), and the predicted probability is a mathematical function of the risk score, as noted above. All of the BPH and cancer cases have predicted probabilities extremely close to 1.0, and all of the normal cases have predicted probabilities of ≤0.002 when BPH or cancer are compared with normal cases. These marked distinctions in predicted probabilities confirm the clear separation of groups, as shown in Fig. 1. When cancer is compared with BPH, predicted cancer probabilities ranged from 0.42 to 1.00 and predicted BPH probabilities ranged from 0.00 to 0.65.

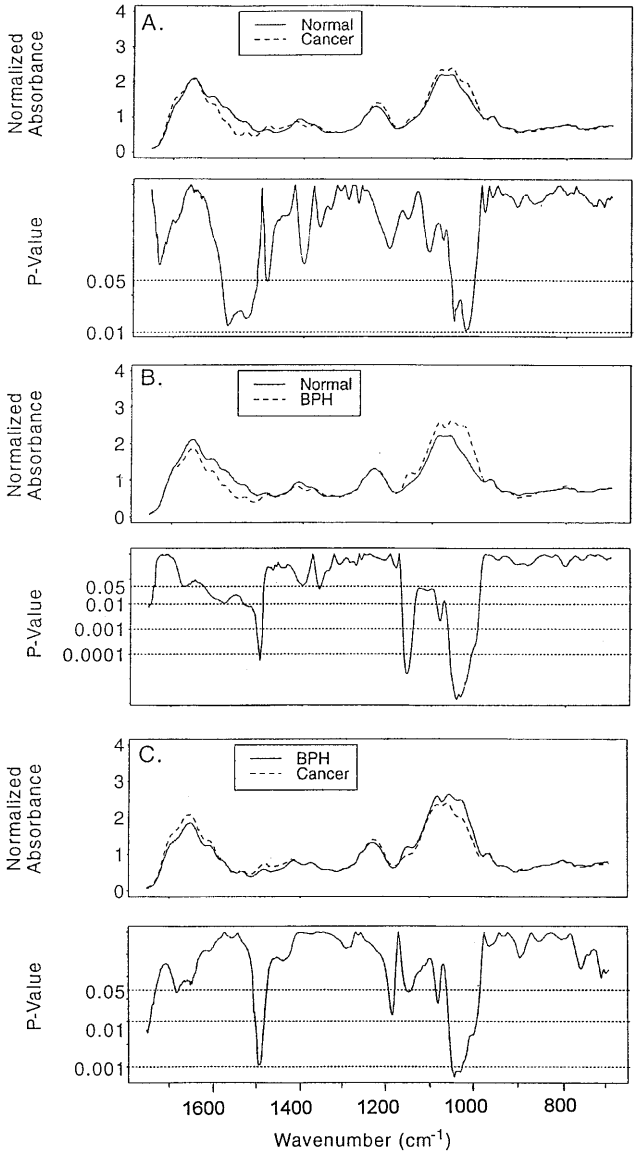

FIG. 2. Comparison of the mean spectrum of cancer vs. normal tissue (*A*), BPH vs. normal tissue (*B*), and cancer vs. BPH (*C*). The lower plot of each panel shows the statistical significance of the difference in mean absorbance at each wavenumber, based on the unequal variance *t* test. *P* values are plotted on the $\log_{10}$ scale.

The two outliers omitted from the analyses tend to support the findings. The outlier BPH and cancer points lie to the right in the PC plot (Fig. 1). This is the same direction found with the progressions from normal to BPH and from normal to cancer, suggesting that the outlier DNAs have a higher degree of structural modification. When the models shown in Table 2 were used to classify the two outliers, the BPH outlier was correctly classified, using the normal vs. BPH model, with a predicted BPH probability close to 1.0. The cancer outlier is also correctly classified in the normal vs. cancer model with a predicted cancer probability close to 1.0. In the BPH vs. cancer model, the BPH outlier is correctly classified with a predicted cancer probability close to zero; however, the cancer outlier is incorrectly classified as a BPH with a cancer probability close to zero.

Age and Gleason Score Relationships. Age does not appear to be a factor in creating the pronounced distinctions among groups, although the incidence of prostate cancer increases significantly over the age of 50 years (1). The age ranges for the three groups were 16–73 years for normal ($n = 5$); BPH, 58–73 ($n = 17$); and cancer, 61–76 ($n = 7$). Among the Spearman correlations of age with each of the four PC scores, none were statistically significant ($P < 0.05$). In all, 28 correlations were considered, consisting of age correlated with each PC score in each of the three groups, as well as in all pairs of groups (e.g., age correlated with each PC score in normal and BPH tissue combined) and in the entire pooled set of 29 cases. Spearman correlations ranged in magnitude from 0.01 to 0.59 with $P = 0.09$ to $P = 1.0$. The most significant correlation was $r = -0.51$ between age and PC4 in the combined normal and cancer groups ($P = 0.09$). When PC4 was omitted from the logistic

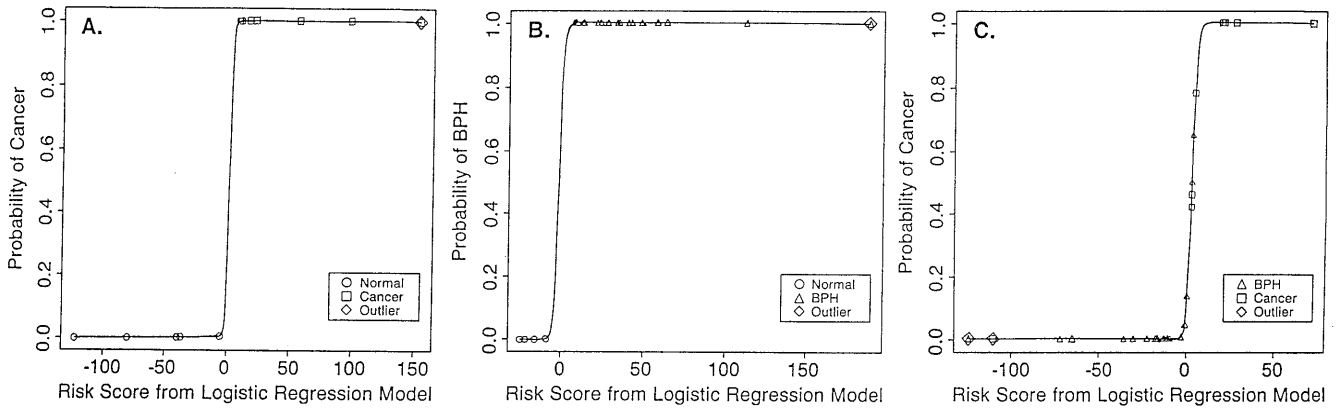

FIG. 3. Sigmoid curves depicting the probability of DNA being classified as normal tissue vs. cancer (*A*), normal tissue vs. BPH (*B*), and BPH vs. cancer (*C*). The curves are based on the logistic regression models depicted in Table 2. The predicted probabilities rise very rapidly over a narrow range, which reflects a high degree of discrimination among groups and a precipitous change in DNA structure associated with the normal to BPH and normal to cancer progressions. Each sample is plotted at its predicted probably.

regression analysis and models were based on PC1–PC3, the *P* values corresponding to those in Table 2 were, from top to bottom, $P < 0.001$, $P < 0.001$, and $P = 0.005$, again supporting a nonrandom distinction among the groups. These results based on PC4 and the weak or nonsignificant correlations between age and other PC scores do not support any role for age in the ability to use spectra to distinguish among the groups. Comparable studies employing the PCA/FT-IR technology to breast tissues (normal and cancer) also revealed only weak correlations of PC scores with patient age (11, 12). Moreover, other studies of DNA from the normal, benign and malignant female breast (24) and the human brain (25) did not show a relationship between 8-hydroxydeoxyguanosine concentrations in tissues and patient age. The present findings, which are consistent with the results from these studies, suggest that the relatively high oxidative burden in tissues is likely a predominant factor in prostate cancer and transcends patient age.

The Gleason score, which uses microscopically evinced architectural changes to classify tumor status (1), had little association with the PC scores, although based on the $n = 7$ cancer cases, there was limited power to detect other than strong associations. Spearman correlations of PC scores 1–4 with the Gleason score ranged from -0.49 to $+0.26$, with $P = 0.2$ to 0.8. Further studies will be required to establish whether the DNA alterations are correlated with changes at the cellular level; however, both the Gleason and PC scores reflect complex suites of underlying biological changes that are not easily identified.

Logistic Regression Models of Probability. The findings suggest that the ·OH likely plays a major role in the transformation of normal prostate tissue to the cancer state, which is consistent with its proposed role in breast cancer (7–12). However, the separation of clusters (Fig. 1) is more pronounced for the prostate than for normal tissue, primary cancer, and metastatic cancer of the breast (11, 12). In the breast cancer studies, logistic regression models of infrared data on DNA from normal and cancer tissue yielded a sensitivity and specificity of about 83%, whereas the comparable prostate data (Table 2) yielded 100% sensitivity and specificity, though with a smaller group of samples.

The sigmoid curves (Fig. 3) for the prostate show sharp transitions between the normal and cancer states and normal and BPH states. These transitions are characterized by a lack of cases at intermediate probabilities, corresponding to the clear separation of groups in Fig. 1. A sharp transition was also obtained with sigmoid curves in which the $\log_{10}$ concentration ratio of ·OH-modified bases was plotted against the probability of breast cancer (9). The sensitivity was 91% and the specificity ranged from 94% to 97%, depending on the modified base model used. Thus, at some point in the modification of DNA, critical structural changes apparently take place that lead to a rapid increase in cancer probability.

BPH is not known to be etiologically related to prostate cancer; however, it is of interest that the BPH vs. prostate cancer curve (Fig. 3*C*) shows several cases having intermediate probabilities. The configuration of cases in Fig. 1 also provides some insight into the controversial view that BPH is a direct precursor of prostate cancer (1). The findings do not support this concept in that the BPH group lies "beyond" the cancer group, starting from the normal group. This positioning suggests that a transition from BPH to cancer would involve a reversal of some of the spectral transitions shown to be associated with cancer, or that there are additional changes in the BPH DNA that mimic a reversal in the progression to cancer. Alternatively, modifications may result in DNA structures that lead to a variety of nonneoplastic lesions, including BPH. Support for this concept comes from studies of English sole exposed to environmental chemicals (18) which showed significant correlations between 8-OH-Ade and 8-OH-Gua and five nonneoplastic hepatic lesions (including the putatively preneoplastic lesion, basophilic foci). The results suggested that the ·OH is likely a common factor in the etiology of both the modified bases and the nonneoplastic lesions. Although BPH may not be a direct precursor of prostate cancer, PCA/FT-IR spectral analysis may provide a promising means of predicting the occurrence of prostate cancer, based on the structural status of BPH DNA.

The absence of transition states in the normal to cancer and normal to BPH curves is of interest. This is likely due to the fact that "transition" tissues having DNA values between zero and 100% probability (Fig. 3) were not part of this study. Clearly, additional research with a larger number of samples is necessary to obtain information on the ability of the PCA/FT-IR technology to detect transition states associated with the normal to BPH and normal to prostate cancer progressions. Additional studies also seem warranted to test the important hypothesis that the PCA/FT-IR spectral analysis of DNA from prostate tissue (≈ 20 μg is required) will provide a sensitive means for screening and predicting prostate cancer.

Evidence with the breast (7–12, 17), and now the prostate, suggests that DNA structure is progressively altered in response to factors in the microenvironment, notably the ·OH, that are likely etiologically related to the development of cellular lesions—breast tumors (primary and metastatic invasive ductal carcinoma) (7–12), prostate tumors (adenocarcinoma), and BPH. It is suggested that intervention to forestall or correct the genetic instability of these tissues and likely

increase in cancer risk should focus on controlling the cellular redox status and ·OH concentrations. The approaches may include control of the iron-catalyzed conversion of H_2O_2 to the ·OH (15), regulation of ·OH production through redox cycling of hormones (26) and environmental xenobiotics (27), and antioxidant/reductant therapy (28, 29).

We thank the Cooperative Human Tissue Network (Cleveland) for providing prostate tissues and pathology data; Dr. Henry Gardner for interest and support; and Derek Stanford for excellent computing. Helpful comments on the manuscript were provided by Drs. Anthony Frank, Joachim Liehr, and Charles Waldren. This work was supported by U.S. Army Medical Research and Materiel Command Contract DAMD17-95-1-5062.

1. Kirby, R. S., Christmas, T. J. & Brawer, M. (1996) *Prostate Cancer* (Mosby, London).
2. Isaacs, W. B., Bova, G. S., Morton, R. A., Bussemakers, M. J. G., Brooks, J. D. & Ewing, C. M. (1995) *Cancer Surv.* **23,** 19–32.
3. Winter, M. L. & Liehr, J. G. (1996) *Toxicol. Appl. Pharmacol.* **136,** 211–219.
4. Halliwell, B. & Aruoma, O. I. (1991) *FEBS Lett.* **281,** 9–19.
5. Feig, D. I., Reid, T. M. & Loeb, L. A. (1994) *Cancer Res.* **54,** 1890s–1894s.
6. Kuchino, Y., Mori, F., Kasai, H., Inoue, H., Iwai, S., Miura, K., Ohtsuka, E. & Nishimura, S. (1987) *Nature (London)* **327,** 77–79.
7. Malins, D. C. & Haiminot, R. (1991) *Cancer Res.* **51,** 5430–5432.
8. Malins, D. C. (1993) *J. Toxicol. Environ. Health* **40,** 247–261.
9. Malins, D. C., Holmes, E. H., Polissar, N. L. & Gunselman, S. J. (1993) *Cancer* **71,** 3036–3043.
10. Malins, D. C., Polissar, N. L., Nishikida, K., Holmes, E. H., Gardner, H. S. & Gunselman, S. J. (1995) *Cancer* **75,** 503–517.
11. Malins, D. C., Polissar, N. L. & Gunselman, S. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 2557–2563.
12. Malins, D. C., Polissar, N. L. & Gunselman, S. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 14047–14052.
13. Kamiya, H., Miura, H., Murata-Kamiya, N., Ishikawa, H., Sakaguchi, T., Inoue, H., Sasaki, T., Masutani, C., Hanaoka, F., Nishimura, S. & Ohtsuka, E. (1995) *Nucleic Acids Res.* **23,** 2893–2899.
14. Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S. & Loeb, L. A. (1992) *J. Biol. Chem.* **267,** 166–172.
15. Imlay, J. A., Chin, S. M. & Linn, S. (1988) *Science* **240,** 640–642.
16. Ward, J. F., Evans, J. W., Limoli, C. L. & Calabro-Jones, P. M. (1987) *Br. J. Cancer* **55,** 105–112.
17. Olinski, R., Zastawny, T. H., Foksinski, M., Barecki, A. & Dizdaroglu, M. (1995) *Free Radical Biol. Med.* **18,** 807–813.
18. Malins, D. C., Polissar, N. L., Garner, M. M. & Gunselman, S. J. (1996) *Cancer Res.* **56,** 5563–5565.
19. Parker, F. S. (1983) *Application of Infrared, Raman, and Resonance Raman Spectroscopy in Biochemistry* (Plenum, New York), pp. 349–398.
20. Tsuboi, M. (1969) *Appl. Spectrosc. Rev.* **3,** 45–90.
21. Tsuboi, M. (1974) in *Basic Principles in Nucleic Acid Chemistry,* eds. Ts'o, P. O. P. (Academic, New York), Vol. 1, pp. 399–452.
22. Von Sonntag, C., Hagen, U., Schön-Bopp, A. & Schulte-Frohlinde, D. (1981) *Adv. Radiat. Biol.* **9,** 110–142.
23. Morgan, W. F., Day, J. P., Kaplan, M. I., McGhee, E. M. & Limoli, C. L. (1996) *Radiat. Res.* **146,** 247–258.
24. Musarrat, J., Arezina-Wilson, J. & Wani, A. A. (1996) *Eur. J. Cancer* **32A,** 1209–1214.
25. Sanchez-Ramos, J. R., Overvik, E. & Ames, B. N. (1994) *Neurodegeneration* **3,** 197–204.
26. Han, X. & Liehr, J. G. (1995) *Carcinogenesis* **16,** 2571–2574.
27. Bagchi, D., Bagchi, M., Hassoun, E. A. & Stohs, S. J. (1995) *Toxicology* **104,** 129–140.
28. Ames, B. N., Shigenaga, M. K. & Hagen, T. M. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7915–7922.
29. Bast, A., Haenen, G. R. M. M. & Doelman, C. J. A. (1991) *Am. J. Med.* **91**, Suppl. 3C, 2S–13S.