Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 7637–7642, June 1998
Medical Sciences

A unified theory of carcinogenesis based on order–disorder transitions in DNA structure as studied in the human ovary and breast

(cancer etiology/hormone responsive tissues/Fourier transform-infrared spectroscopy/chaos theory/free radicals)

DONALD C. MALINS*†, NAYAK L. POLISSAR‡, STEFAN SCHAEFER*, YINGZHONG SU*, AND MARK VINSON*

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

Contributed by Donald C. Malins, April 29, 1998

ABSTRACT **Fourier transform-infrared/statistics models demonstrate that the malignant transformation of morphologically normal human ovarian and breast tissues involves the creation of a high degree of structural modification (disorder) in DNA, before restoration of order in distant metastases. Order–disorder transitions were revealed by methods including principal components analysis of infrared spectra in which DNA samples were represented by points in two-dimensional space. Differences between the geometric sizes of clusters of points and between their locations revealed the magnitude of the order–disorder transitions. Infrared spectra provided evidence for the types of structural changes involved. Normal ovarian DNAs formed a tight cluster comparable to that of normal human blood leukocytes. The DNAs of ovarian primary carcinomas, including those that had given rise to metastases, had a high degree of disorder, whereas the DNAs of distant metastases from ovarian carcinomas were relatively ordered. However, the spectra of the metastases were more diverse than those of normal ovarian DNAs in regions assigned to base vibrations, implying increased genetic changes. DNAs of normal female breasts were substantially disordered (e.g., compared with the human blood leukocytes) as were those of the primary carcinomas, whether or not they had metastasized. The DNAs of distant breast cancer metastases were relatively ordered. These findings evoke a unified theory of carcinogenesis in which the creation of disorder in the DNA structure is an obligatory process followed by the selection of ordered, mutated DNA forms that ultimately give rise to metastases.**

The hydroxyl radical (•OH) is a highly reactive chemical species that is known to alter the structure of DNA in human tissues (1–4). In hormone responsive tissues (e.g., the human breast), this radical is believed to be produced via the metal (e.g., Fe^{2+})-catalyzed decomposition of H_2O_2, which may arise from redox cycling of catecholestrogen metabolites (5) and xenobiotics (e.g., aromatic hydrocarbons) (6, 7). The •OH-induced modification of DNA involves the nucleotide base and phosphodiester-deoxyribose structures (3, 8). The reaction rate with the bases has been estimated to be approximately four times that occurring with deoxyribose (8). The base reactions produce mutagenic derivatives, such as 8-hydroxyguanine (8-OH-Gua) and 8-hydroxyadenine (8-OH-Ade), together with the putatively nonmutagenic ring-opened structures 2,6-diamino-4-hydroxy-5-formamidopyrimidine (Fapy-G) and 4,6-diamino-5-formamidopyrimidine (Fapy-A) (2, 9). The •OH also abstracts hydrogen atoms from the furanose ring of deoxyribose (8), which produces a variety of hydroxy derivatives that lead to strand breaks and the loss of phosphoric acid (8). Accordingly, the attack of the •OH creates substantial disorder likely reflected in the formation of potentially billions of new DNA structures (10), some of which may give rise to altered protein expression and function.

Fourier transform-infrared (FT-IR)/statistics models were used to demonstrate cancer-related structural alterations in the DNA of human tissues (3, 11–12). The models, using principal components analysis, allowed a group of spectra to be represented as points in space, each point being a highly discriminating measure of DNA structure. In the human female breast, the DNA modifications (e.g., free radical-induced) produce alterations in the vibrational and rotational motion of functional groups (hence the FT-IR spectra), thus shifting the location of the points. Different clusters of points had different sizes, shapes, and locations in principal components (PC) plots, depending on whether the DNAs were from normal breast tissue, a primary cancer, or a metastasized primary cancer (a primary cancer that has disseminated metastases) (10, 11). These changes reflected various degrees of structural disorder (chaos) in DNA. The disorder was linked to tumor formation and constituted a basis for constructing cancer probability relationships with a high sensitivity and specificity by using logistic regression models. The models related the FT-IR wavenumber–absorbance spectra (including differences between absorbancies of base and phosphodiester-deoxyribose structures) to groups of tissues, as well as assigning probability (e.g., for cancer) to individual tissues (3, 10, 11).

FT-IR/statistics models also were used to describe the progression of the morphologically normal prostate to prostatic adenocarcinoma and benign prostatic hyperplasia (13). Significant differences were evident between the mean spectra of the normal prostate and prostatic adenocarcinoma and between the normal prostate and benign prostatic hyperplasia. The clusters of points, as represented by their centroids, had different locations in PC plots (i.e., different disordered states). As with the breast, structural differences in the DNAs between groups of tissues were the basis for constructing cancer/benign prostatic hyperplasia probability vs. risk score relationships having a high degree of sensitivity and specificity.

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.

Abbreviations: PC, principal component; FT-IR, Fourier transform-infrared; •OH, hydroxyl radical; 8-OH-Gua, 8-hydroxyguanine; 8-OH-Ade, 8-hydroxyadenine; Fapy-G, 2,6-diamino-4-hydroxy-5-formamidopyrimidine; Fapy-Ade, 4,6-diamino-5-formamidopyrimidine; O_n, DNA of morphologically normal ovarian tissue; AC, DNA of primary ovarian adenocarcinoma; AC_m, DNA of metastasized primary ovarian adenocarcinoma; AC_{dm}, DNA of ovarian metastases; RMT, DNA of reduction mammoplasty tissue; IDC, DNA of invasive ductal carcinoma; IDC_m, DNA of metastasized invasive ductal carcinoma; IDC_{dm}, DNA of invasive primary ductal carcinoma metastases; HBL, DNA of human blood leukocytes; HNT, DNA of hypothetically normal tissue.
†After the first author, the authors are listed in alphabetical order.

The studies described led to the concept that the creation of structural disorder in DNA is a dynamic process intimately involved in cellular transformations (e.g., of normal cells to the malignant state). Using the powerful FT-IR/statistics technology, the present studies of the human ovary, breast, and their corresponding cancers test the hypothesis that the creation of disorder (represented by PC cluster diversity and location) in DNA structure is a critical factor in the transformation of morphologically normal tissues into primary and metastatic tumors.

MATERIALS AND METHODS

Tissue Acquisition and DNA Isolation. Samples of human ovarian tissues and breast metastases were obtained from the National Cancer Institute Cooperative Human Tissue Network. Other breast samples were acquired previously (2, 3). The ovary and breast samples were obtained from women with a mean age of 56 ± 16 and 51 ± 21 years, respectively. No extraneous histologies were evident in any of these tissues. Evidence for metastasized primary tumors was based on the identification of metastases at distant sites. DNA was isolated from each sample and purity was established spectroscopically (2). DNA from the ovaries was obtained from 13 morphologically normal tissues (O_n), six primary adenocarcinomas (AC), nine metastasized primary adenocarcinomas (AC_m), and seven distant metastases to the colon (AC_{dm}). DNA from the breast was obtained from 19 reduction mammoplasty tissues (RMT) of patients who had undergone hypermastia surgery, 10 invasive ductal carcinomas (IDC), 23 metastasized IDCs (IDC_m), and 7 samples of distant metastases to axillary nodes (IDC_{dm}). DNA from human blood leukocytes was obtained from five healthy individuals.

FT-IR/Statistical Analysis. This analysis was carried out primarily as described (3, 10–12). In brief, the procedure involves the use of a FT-IR microscope spectrometer. A thin film of DNA is placed on a BaF_2 window, and an IR beam is focused on it. The interferogram recorded in the detector is then Fourier-transformed into an absorbance spectrum that is baselined and normalized to an absorbance of 1.0 in the range of interest (e.g., 1,750 to 770 cm^{-1}). For wavenumber-by-wavenumber analyses (e.g., *t* tests), spectra were split into two regions (1,750 to 1,350 cm^{-1} and 1,314 to 770 cm^{-1}) and independently normalized. To develop a common basis for plotting, PC scores for the entire sample database (ovarian and breast tissues) were calculated giving equal weight to each group. The difference between two DNA spectra or between the centroids (mean spectra) of two groups was defined as the Euclidean distance. This distance was expressed as a percentage by dividing it by the square root of the number of wavelengths (i.e., 1,750 to 770 cm^{-1}), then dividing by the mean normalized absorbance and multiplying by 100. We used the permutation test (5×10^2 permutations) to test the null hypothesis that the distance between centroids of two groups is 0 (i.e., that the mean spectra are the same for the two groups) and that the observed distance arises by chance. We also used a two-sided unequal variance *t* test for the null hypothesis that the mean absorbance at a given wavenumber is equal between groups. The *t* test, carried out at each wavenumber, yields a plot of *P* values vs. wavenumber. Re-sampling (with 10^3 samples) was used to test the null hypothesis that the distance between states (e.g., between the centroid for normal tissue and that for primary tumor tissue) is the same for the ovary and breast. The same re-sampling procedure was used to compare the "base" region (1,750 to 1,315 cm^{-1}) to the "phosphodiester-deoxyribose" region (1,314 to 770 cm^{-1}). The *P* value for these re-sampling tests is defined as twice the proportion of re-sampled observations that are on the opposite side of 0 from the observed differences, with a maximum of $P = 1.0$. We tested for differences in diversity between two groups, wavenumber-by-wavenumber, based on the ratio of group variances at each wavenumber by using a two-sided F test. Differences in PC cluster size and/or location were determined by using a test for the equality of covariance matrices of PC scores 2–6. Regression analysis of age vs. PC scores 2–6 was used to test whether age played a role in determining spectral characteristics. All hypothesis testing and plotting was carried out by using the SAS and S-PLUS statistical packages.

RESULTS

Change in mean distance from the centroid (diversity) and/or change in mean spectra (PC location) are both measures of alterations in the order–disorder status of cellular DNA. As an example, differences between mean spectra of the O_n and AC groups are illustrated in Fig. 1*A*. $P \leq 0.05$ delineates wavenumber regions in which the more significant differences exist (Fig. 1*B*). These differences also are consistent with the substantial change in centroid location between the O_n and AC groups (Table 1). A significant change was not found between the mean spectra of the AC and AC_m groups; however, a significant spectral change was evident in the transition from the AC_m to the AC_{dm}. There was not a significant difference between the mean spectra of the O_n and the AC_{dm}. The O_n was a relatively tight group, whereas the AC was highly diverse. The

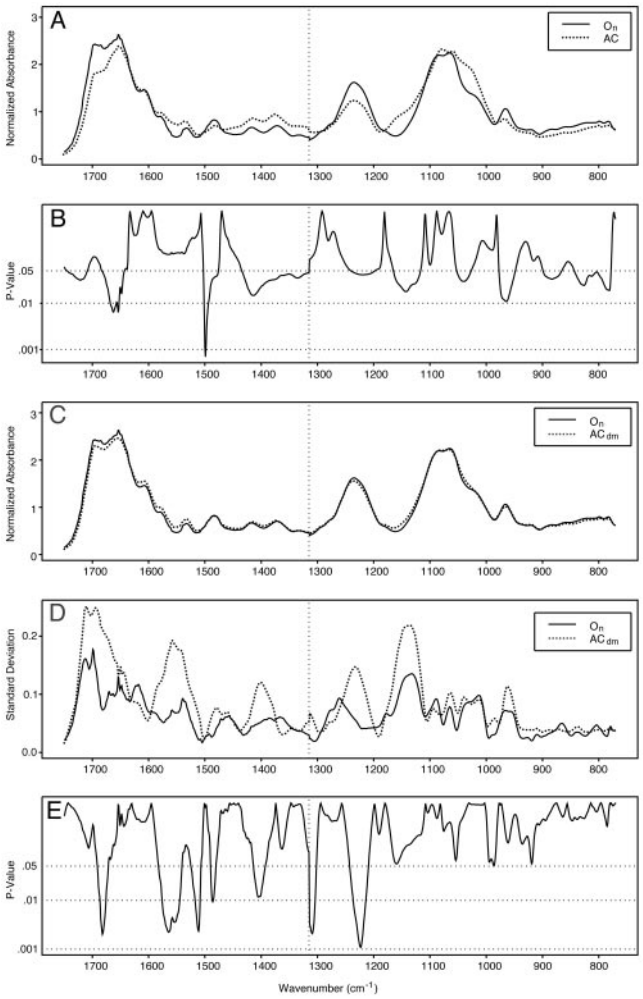

FIG. 1. DNA spectral comparisons: (*A*) grand mean spectra of morphologically normal ovarian tissue (O_n) and primary ovarian adenocarcinoma (AC); (*B*) *P* values for spectral comparison in *A*; (*C*) grand mean spectra of O_n with AC metastases to colon (AC_{dm}); (*D*) SDs between spectral comparisons in *C*; and (*E*) *P* values for SD comparisons for *D*.

Table 1. Order–disorder comparisons between DNA structures in the neoplastic transformation of morphologically normal breast and ovarian tissues

Groups compared		Difference between grand mean spectra as percentage		Diversity: mean difference from group mean spectrum as percentage		
1	2	Percentage	*P* value*	Group 1: mean ± SD	Group 2: mean ± SD	*P* value†
Ovary						
O_n	AC	23.5	<0.002	8 ± 3	20 ± 8	0.001
AC	AC_m	6.2	0.7	20 ± 8	16 ± 6	0.4
AC_m	AC_{dm}	16.1	0.02	16 ± 6	9 ± 4	0.008
O_n	AC_{dm}	6.2	0.1	8 ± 3	9 ± 4	0.4
Breast						
RMT	IDC	9.3	<0.002	10 ± 5	8 ± 3	0.4
IDC	IDC_m	7.0	0.09	8 ± 3	13 ± 5	0.003
IDC_m	IDC_{dm}	16.1	0.002	13 ± 5	10 ± 4	0.1
RMT	IDC_{dm}	16.3	<0.002	10 ± 5	10 ± 4	0.9

*One sided *P* values based on permutation test.
†Two sided *P* values based on Mann–Whitney test.

diversities of the AC and AC_m groups were similar, and the AC_{dm} was a substantially tighter cluster than the relatively diverse AC_m group. Table 1 shows that the O_n and the AC_{dm} had similar diversities and comparable mean spectra. However, Fig. 1 *D–E* shows that the O_n and the AC_{dm} groups had substantially different patterns of diversity (different SDs; Fig. 1*D*) at a number of wavelengths, particularly in the left area of the spectrum (base vibrations above ≈1,315 cm^{-1}). Many of these differences in diversity yielded $P \leq 0.05$ (Fig. 1*E*). The null hypothesis that the two groups have the same diversity pattern (identical wavenumber-by-wavenumber SDs across the spectrum) is rejected with $P = 0.02$, based on the covariance matrices for PC scores 2–6.

The differences in mean spectra and diversity (Table 1) of the $O_n \rightarrow AC$, $AC \rightarrow AC_m$, and $AC_m \rightarrow AC_{dm}$ transitions are illustrated graphically in PC plots using the second and third PC scores (Fig. 2). The differences in diversity and locations of the O_n and AC clusters are evident in Fig. 2*A* in which the AC cluster is more diverse and shifted to the left of the O_n. In Fig. 2*B*, the AC and AC_m clusters occupy about the same PC location and are equally diverse, reflecting their similar mean spectra and diversity (Table 1). Fig. 2*C* shows that the AC_m and the AC_{dm} differ both in location and diversity. In Fig. 2*D*, the AC_{dm} and O_n samples overlap considerably and are about equal in size, each representing a tight cluster. However, as indicated, the two groups differ in their SDs at certain wavenumbers (Fig. 1 *D–E*). That is, disorder in different DNA structures (as represented by the spectral properties) distinguishes the two groups.

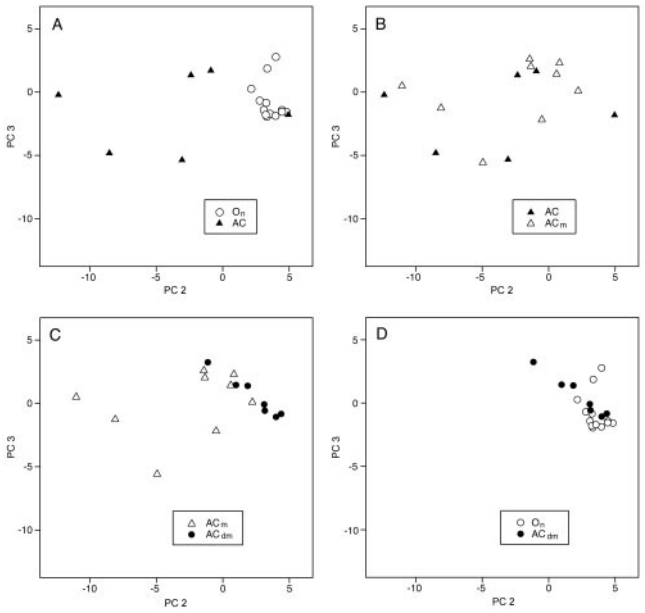

FIG. 2. PC plots comparing spectra of ovarian DNAs from (*A*) morphologically normal tissue (O_n) and primary adenocarcinoma (AC); (*B*) AC and metastasized primary AC (AC_m); (*C*) AC_m with AC metastases to the colon (AC_{dm}); and (*D*) a comparison of O_n with AC_{dm}. See text and Table 1 for statistical comparisons of order–disorder status between groups.

Cluster analysis showed that the various stages of ovarian tumor progression comprise a mixture of subgroups (i.e., disorder mixed with relative order). Fig. 3 shows a cluster analysis of FT-IR spectra depicting the Euclidean distance, expressed as a percentage, between each spectrum and its "nearest neighbor." The O_n shows a fairly tight cluster with no nearest neighbor distances beyond about 10% (Fig. 3*A*). The AC cluster (Fig. 3*B*) shows a wide range between spectra, with some as close as 6% and some as distant as 30%. The AC_m cluster (Fig. 3*C*) appears to be a mixture between a tight subgroup of DNAs (lower in the panel) that have no more than

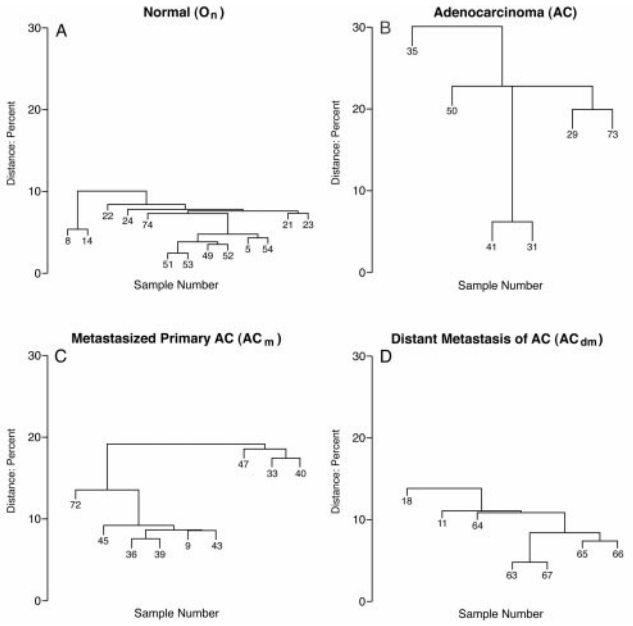

FIG. 3. Cluster analysis of ovarian spectra of DNAs. This analysis is based on the distance of each sample to its nearest neighbor. The *y* axis shows the percentage difference between spectra (e.g., 31 is ≈6% different from 41 in *B*).

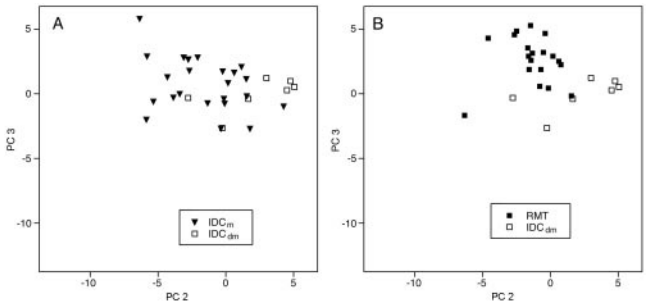

FIG. 4. PC plots comparing spectra of human breast DNAs from (*A*) metastasized primary adenocarcinoma (IDC$_m$) and IDC metastases to axillary nodes (IDC$_m$); and (*B*) morphologically normal reduction mammoplasty tissue (RMT) and IDC$_{dm}$. See text and Table 1 for statistical comparisons of order–disorder status between groups.

a 10% nearest neighbor distance and a second relatively diverse subgroup (higher in the panel) with 18–19% nearest neighbor distance. All spectra in the second subgroup are at least 25% distant from spectra in the first subgroup. An individual spectrum (sample 72) appears at an intermediate distance from the two subgroups. The AC$_{dm}$ group (Fig. 3*D*) is relatively tight, with nearest neighbor distances not exceeding 14%.

The breast samples also show substantial differences between groups both in PC cluster diversity and mean spectra, although the range of differences in mean spectra between groups and the range of diversities are not as great as among the ovary samples (Table 1). The differences between groups range from 9 to 16% for the breast samples (compared with 6–24% for the ovary samples), and the mean distance from the centroid varies from 8 to 13% (compared with 8–20% for the ovary). The transition RMT → IDC is characterized by a moderate, but highly significant, difference between mean spectra; however, there was not a significant change in the mean difference from the centroid. The transition IDC → IDC$_m$ shows a marginally significant mean change but a notably significant change in diversity. The difference in the mean spectra of IDC$_m$ → IDC$_{dm}$ is also significant, with no significant change in diversity. The IDC$_{dm}$, the terminal stage in the sequence of transitions, has a significantly different mean spectrum from the RMT with very little difference in diversity. Comparisons between the RMT, IDC$_m$, and IDC$_{dm}$ clusters are shown in the PC plots of Fig. 4. Fig. 4*A* shows a substantial overlap between the IDC$_m$ and the IDC$_{dm}$ clusters; however, the IDC$_{dm}$ cluster is more compact, representing a more ordered state as reflected by its smaller mean distance to the centroid. Fig. 4*B* shows little overlap between the RMT and IDC$_{dm}$ clusters, indicating different mean spectra.

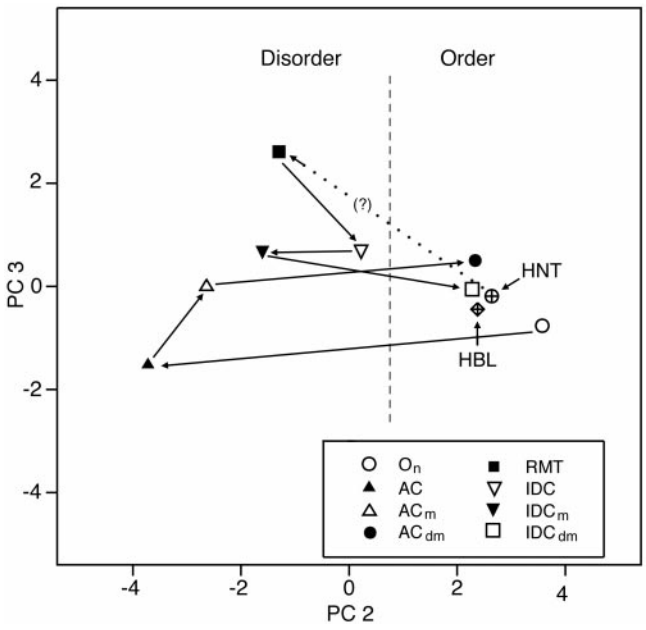

FIG. 5. Tumor progression pathways are depicted in a PC plot of human ovarian and breast DNA centroids (derived from groups of spectra). The centroid representing the DNAs of hypothetically normal tissue (HNT) and human blood leukocytes (HBL) also are included (see text for details). The dashed vertical line broadly distinguishes the centroids of the relatively ordered and disordered groups.

The sequence of ovary and breast DNA transitions is illustrated in the PC plot of Fig. 5, which depicts the centroids of each cluster. The ovary DNAs show a substantial "leap" from the centroid of the O$_n$ group (in the "order" region) to the AC centroid, a short step back to the AC$_m$ centroid, and finally a shift to the AC$_{dm}$ centroid located close to that of the O$_n$ (6% distance). The breast centroids proceed along a different path but ultimately converge on the order region, as occurs with the ovary (Fig. 5). The RMT, IDC, and IDC$_m$ centroids are located in the "disorder" region. The final stage of the progression, represented by the IDC$_{dm}$ centroid, is close to that of the O$_n$, the AC$_{dm}$, and the HBL centroids. Also included is a hypothetical normal tissue DNA (HNT) centroid, which is the mean of the O$_n$, AC$_{dm}$, IDC$_{dm}$ and HBL centroids. The centroid of the HNT is intended to serve as a reference point and speculative origin for essentially unmodified normal breast DNAs.

Comparisons of changes in the base (1,750 to 1,315 cm^{-1}) and phosphodiester-deoxyribose (1,314 to 770 cm^{-1}) regions are given in Table 2, plus the sum of transitions for both of

Table 2. Length of centroid-to-centroid transition pathways, as percent, for ovary and breast carcinogenesis

	1,750 to 1,315 cm^{-1} Base region		1,314 to 770 cm^{-1} Deoxyribose region		Difference Base–deoxyribose regions		
Transition	Distance, %	SE*	Distance, %	SE*	Difference	SE*	*P* value*
Ovary							
O$_n$ to AC	24.1	8.3	20.8†	6.5	3.3	4.6	0.5
AC to AC$_m$	3.9	5.3	4.4	4.7	−0.5	4.4	1.0
AC$_m$ to AC$_{dm}$	17.8	5.7	13.3	4.2	4.5	3.5	0.2
Total path	45.8	11.7	38.5	10.4	7.3	7.5	0.2
Breast							
RMT to IDC	11.9	2.2	6.2†	1.3	5.7	1.8	<0.001
IDC to IDC$_m$	8.7	3.0	4.4	1.5	4.3	2.0	0.02
IDC$_m$ to IDC$_{dm}$	21.2	4.4	10.0	1.9	11.2	3.0	0.01
Total path	41.8	5.7	20.6	2.7	21.2	4.9	<0.001

*From resampling.
†P = 0.03 for ovary transition distance compared with corresponding breast transition distance; based on resampling.

these spectral regions in relation to the breast and ovary DNAs. On the basis of the percentage change in spectra between states, the total cancer progression involves remarkably large structural changes in DNA, as indicated by the total path length of 46% for the ovarian base region and 39% for the corresponding phosphodiester-deoxyribose region (Table 2). Comparable results for the breast were also substantial: 42% for the base region and 21% for the phosphodiester-deoxyribose region. Considering that the RMT is reported to be significantly modified (2, 3), the path length between the possible starting point of the breast cancer process (the HNT centroid) and the RMT centroid (Fig. 5) contributes substantially to the total path distance of the breast. This distance was 25% for the base region and 9% for the phosphodiester-deoxyribose region. The total path length, if the HNT to RMT path were included, would be 67% for the base region and 30% for the phosphodiester-deoxyribose region. Patient age appears to play a negligible role in determining spectral differences between the ovarian and breast groups. This result was established on the basis of regression analyses of age in relation to PC scores.

DISCUSSION

In previous studies (3, 10–12), FT-IR/statistics models provided the first evidence showing that cellular transformations in breast tumor formation (e.g., RMT $\rightarrow$ IDC $\rightarrow$ IDC_m) involve order–disorder transitions in DNA structure. Subsequent studies that were performed using these models (13) showed that the transformation of morphologically normal prostate [(normal $\rightarrow$ adenocarcinoma) and (normal $\rightarrow$ benign prostatic hyperplasia)] also produces discernible changes in the order–disorder status of DNA. These initial studies raised the important question of whether changes in DNA, as determined by using FT-IR/statistics, represent critical events on which cancer progression depends to reach the stage of distant metastases.

In the ovary, the transition $O_n \rightarrow AC$ represents a major change in the DNAs from a relatively ordered to a substantially disordered state (Table 1; Figs. 2*A* and 5). Pronounced alterations in areas of the spectra assigned to both base and phosphodiester-deoxyribose structures (Table 2) reflected the global nature of these alterations. The order–disorder status was virtually unchanged in the transition $AC \rightarrow AC_m$ (Table 1; Fig. 2*B*); however, the transition to the AC_{dm} resulted in a major change toward the reinstatement of order, comparable to that of the O_n, as indicated by the differences in mean spectra and cluster diversities (Table 1; Fig. 1 *C–D*). The data on SDs of spectra (Fig. 1 *D–E*) further demonstrated that, despite the comparable mean spectra of the O_n and AC_{dm}, differences exist between these groups in vibrations associated with the base and phosphodiester structures. This is consistent with the presence of abundant mutations that characterize metastases (14). Moreover, the highly significant difference in the PO_2^- structure ($\approx$1,250 cm^{-1}) (Fig. 1*E*) likely arose from alterations in base pairing, which would be expected to disrupt the arrangement of the phosphate groups along the DNA backbone, thus altering the vibrational properties of the PO_2^- group. A comparable analysis of the breast data was not appropriate because the RMT samples were disordered.

Cluster analysis (Fig. 3) provided additional insight into the nature of the changes in the ovarian DNAs, showing, for example, that the disordered AC (Fig. 3*B*) and AC_m (Fig. 3*C*) each comprise a mixture of subgroups. Of interest is the appearance of a subgroup within the AC_m (samples 47, 33, and 40) that may represent remnants of the AC group and another subgroup (samples 9–72) that exhibits a relatively ordered state similar to that of the AC_{dm} (Fig. 3*D*). These data, together with those in Fig. 1*E*, support the hypothesis that there is selection of ordered, mutated DNAs for the next stage of cancer progression (AC_{dm}) arising from the pronounced degree of disorder found in the AC_m. The magnitudes of the order–disorder transitions in the ovarian DNAs are substantial, as indicated by the path length data (Table 2): 46% for the base region and 39% for the phosphodiester-deoxyribose region. We suggest that the great number of different DNAs produced in these transitions provide a pool from which viable molecular structures can be selected, consistent with the ultimate attainment of metastases.

In the breast, the creation of disordered DNAs in the transitions RMT $\rightarrow$ IDC $\rightarrow$ IDC_m was reported previously (3, 10–12). The inclusion of data on the IDC_{dm} in the present study afforded the opportunity to explore the nature of tumor progression in the breast to the stage of axillary node metastases. The disorder in the RMT (2, 3) contrasts with the relatively ordered status of the O_n (Table 1; Fig. 5). The magnitude of RMT disorder is substantial based on the path length between this group and the HNT. The 19 RMT samples analyzed had a location distinct from that of the ordered forms, such as the O_n and the HBL, whose centroids are shown on the right side of Fig. 5. A possible explanation for the difference in disorder is that the morphologically normal breast is under greater oxidative stress than the ovary (e.g., from $^{\bullet}OH$), notably due to estrogen metabolism (5, 17, 18). Previous studies have shown that substantial base modifications exist in normal breast DNAs (2, 3), reaching as high as one base modification in 10^3 normal bases (2). We are unaware of comparable data on the ovary.

The RMT $\rightarrow$ IDC transition involves structural changes (disorder) as reflected in a significant distance between the centroids (Table 1; Fig. 5), without a significant change in cluster diversity. Order–disorder transitions of this type may be mostly intramolecular, involving vertical base residue stacking interactions, for example, that are known to produce significant changes in DNA spectra (19, 20). The IDC $\rightarrow$ IDC_m transition involves a substantial increase in diversity (in contrast to the AC $\rightarrow$ AC_m transition) (Table 1). The IDC_m $\rightarrow$ IDC_{dm} transition exhibited a major shift toward order (Table 1), as shown by the fact that the IDC_{dm} cluster was spatially close to that of the HBL, O_n, and AC_{dm} clusters (Fig. 5). The diversities of the RMT and IDC_{dm} clusters are similar; however, they have different PC locations (Table 1; Fig. 5). The initially formed IDC would be expected to be relatively ordered before being progressively damaged by microenvironmental factors, such as $^{\bullet}OH$, that may be produced from H_2O_2 reported to be "constitutively" generated in primary cancer cells (21). In the developing tumor, the damaged forms of DNA would obscure the detection of the initially formed DNA structures. In this context, the progression of morphologically normal breast tissue to distant metastases may not be fundamentally different from that of the comparable ovarian progression, assuming that the disordered RMT was produced from ordered DNAs (e.g., HNT) at some earlier stage in life, possibly shortly after puberty. We recognize the possibility that ordered breast DNAs may exist in certain human populations, notably those from Asia that have a low incidence of breast cancer (22). The virtual lack of relationship between patient age and the present results is consistent with previous studies of radical-induced changes in DNA of human tissues (10, 11, 23, 24).

The transition from disorder in primary tumors, whether or not they had metastasized, to order in distant metastases may involve a significant change in the cellular redox status of the DNA. Prior studies of the IDC $\rightarrow$ IDC_m transition (11) suggested that a shift toward reductive conditions takes place in metastasized primary breast tumors. The evidence was based on a change in the model $\log_{10}$ (Fapy Ade/8-OH-Ade) reflecting an increase in Fapy Ade as the size of the metastasized primary tumor increased [Fapy derivatives are reported to be preferentially synthesized under reductive con-

ditions (15)]. An additional factor consistent with this apparent shift in redox status is the reported development of hypoxia in transformed tissues (25). The proposed shift toward reductive conditions in the metastasized primary tumor cells would be expected to suppress the progression of oxidative DNA damage, thus helping to preserve (stabilize) DNA structures that ultimately become part of the ordered IDC_{dm} group.

The vertical transfer of electrons from base to base along the helix has been reported to extend to 25 bp so that a structural change at one point would likely trigger structural changes far afield (15). Recent evidence for the long range oxidative repair of thymine dimers further demonstrates this unique property of DNA (16). The overall structure of some forms of DNA (e.g., resulting from disrupted base stacking) in a disordered system could alter protein expression and function well beyond changes associated with the coded information inherent in the linear sequence of bases.

The creation of disorder in DNA out of a relatively ordered system and the ultimate restoration of order may be regarded as a prime example of chaos theory (26). A salient feature of complex biological systems is that chaos created at one level of activity can give rise to order at another level: that is, order arises out of chaos and certain dynamic factors are responsible for its emergence (deterministic chaos) (26). In most complex biological systems, the dynamic processes are elusive; however, several factors may be influential in the present order–disorder transitions. These include the reported preferential attack of the $^{\bullet}OH$ on the base structures compared with the attack on deoxyribose (yielding DNA forms with mutated bases and intact deoxyribose moieties) (8) and the preference shown in DNA polymerization for intact substrates (27). Regardless of the processes involved, it is reasonable to assume that the creation of disorder, before the attainment of order in the DNAs of metastases, is pivotal in tumor development. We find no inconsistency between prior findings relating mutations in growth-controlling genes, such as proto-oncogenes and tumor suppressor genes, to carcinogenesis (28) because the creation of disorder in DNA would be expected to lead to a large number of genetic changes that would increase cancer risk.

The attenuation of the disordered status of DNA through intervention is an attractive possibility for reducing cancer risk. This might be accomplished by using therapeutic agents that reduce cellular $^{\bullet}OH$ concentrations or through diets rich in antioxidants (29). Alternatively, the possibility exists to increase the severity of DNA damage in tumor tissues by using DNA-cleaving molecules having selective anti-cancer activity (30, 31).

In conclusion, the present findings evoke a unified theory of carcinogenesis in which order–disorder transitions in DNA structure at various stages of tumor development result in the selection of ordered, mutated DNA forms including those that ultimately give rise to metastases.

We thank the National Cancer Institute Cooperative Human Tissue Network for providing tissues and pathology data; Dr. Henry S. Gardner for interest and support; and Derek Stanford for statistical computing. Helpful comments were provided by Drs. David L. Eaton, Ingegerd Hellström, Karl E. Hellström, Gary K. Ostrander, Peter D. Senter, and James S. Woods. This work was supported by U.S. Army Medical Research and Materiel Command Contract DAMD17-95-1-5062.

1. Olinski, R., Zastawny, T., Budzbon, J., Skokowski, J., Zegarski, W. & Dizdaroglu, M. (1992) *FEBS Lett.* **309,** 193–198.
2. Malins, D. C., Holmes, E. H., Polissar, N. L. & Gunselman, S. J. (1993) *Cancer* **71,** 3036–3043.
3. Malins, D. C., Polissar, N. L., Nishikida, K., Holmes, E. H., Gardner, H. S. & Gunselman, S. J. (1995) *Cancer* **75,** 503–517.
4. Malins, D. C., Polissar, N. L. & Gunselman, S. J. (1997) *Proc. Natl. Acad. Sci. USA* **94***,* 3611–3615.
5. Liehr, J. G. (1997) *Environ. Health Perspect.* **105,** 565–569.
6. Cavalieri, E. L., Stack, D. E., Devanesan, P. D., Todorovic, R., Dwivedy, I., Higginbotham, S., Johansson, S. L., Patil, K. D., Gross, M. L., Gooden, J. K., *et al.* (1997) *Proc. Natl. Acad. Sci. USA* **94,** 10937–10942.
7. Frenkel, K., Wei, L. & Wei H. (1995) *Free Radical. Biol. Med.* **19,** 373–380.
8. von Sonntag, C., Hagen, U., Schon-Bopp, A. & Schulte-Frohlinde, D. (1981) *Adv. Radiat. Biol.* **9,** 110–142.
9. Dizdaroglu, M. & Gajewski, E. (1990) *Methods Enzymol.* **186,** 530–544.
10. Malins, D. C., Polissar, N. L. & Gunselman, S. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 14047–14052.
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., Su, Y., Gardner, H. S. & Gunselman, S. J. (1997) *Nat. Med.* **3,** 927–930.
13. Malins, D. C., Polissar, N. L. & Gunselman, S. J. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 259–264.
14. Fidler, I. J. & Nicolson, G. L. (1991) in *The Breast: Comprehensive Management of Benign and Malignant Diseases,* eds. Bland, K. I. & Copeland, E. M., III (Saunders, Philadelphia), pp. 262–291.
15. Steenken, S. (1989) *Chem. Rev.* **89,** 503–520.
16. Dandliker, P. J., Holmlin, R. E. & Barton, J. K. (1997) *Science* **275**, 1465–1468.
17. Yager, J. G. & Liehr, J. G. (1996) *Annu. Rev. Pharmacol. Toxicol.* **36,** 203–232.
18. Liehr J. G. (1997) *Eur. J. Cancer Prev.* **6,** 3–10.
19. Tsuboi, M. (1969) *Appl. Spectrosc. Rev.* **3,** 45–90.
20. Tsuboi, M. (1974) *in Basic Principles in Nucleic Acid Chemistry*, ed. Ts'o, P. O. P. (Academic, New York), pp. 399–452.
21. O'Donnell-Tormey J., De Boer, C. J. & Nathan, C. F. (1985) *J. Clin. Invest.* **76,** 80–86.
22. Armstrong, B. & Doll, R. (1975) *Int. J. Cancer* **15,** 617–631.
23. Musarrat, J., Arezina-Wilson, J. & Wani, A. A. (1996) *Eur. J. Cancer* **32,** 1209–1214.
24. Sanchez-Ramos, J. R., Overvik, E. & Ames, B. N. (1994) *Neurodegeneration* **3,** 197–204.
25. Höckel, M., Schlenger, K., Aral, B., Mitze, M., Schäffer, U. & Vaupel, P. (1996) *Cancer Res.* **56,** 4509–4515.
26. Kauffman, S. A. (1993) *The Origins of Order, Self-Organization and Selection in Evolution* (Oxford Univ. Press, New York), pp. 33–67 and 175–182.
27. Joyce, C. M. (1997) *Proc. Natl. Acad. Sci. USA* **94** 1619–1622.
28. Cooper, G. M. (1995) *Oncogenes* (Jones and Bartlett, Boston), pp. 67–177.
29. Schwartz, J. L. (1996) *J. Nutrition* **126,** 1221S–1227S.
30. Hiramoto, K., Fujino, T. & Kikugawa, K. (1996) *Mutat. Res.* **360,** 95–100.
31. Quinlan, G. J. & Gutteridge, J. M. C. (1988) *Free Radicals Biol. Med.* **5,** 341–348.