

## Nanocytology of Rectal Colonocytes to Assess Risk of Colon Cancer Based on Field Cancerization

Dhwanil Damania<sup>1</sup>, Hemant K. Roy<sup>2</sup>, Hariharan Subramanian<sup>1</sup>, David S. Weinberg<sup>3</sup>, Douglas K. Rex<sup>4</sup>, Michael J. Goldberg<sup>2</sup>, Joseph Muldoon<sup>2</sup>, Lusik Cherkezyan<sup>1</sup>, Yuanjia Zhu<sup>1</sup>, Laura K. Bianchi<sup>2</sup>, Dhiren Shah<sup>2</sup>, Prabhakar Pradhan<sup>1</sup>, Monica Borkar<sup>2</sup>, Henry Lynch<sup>5</sup>, and Vadim Backman<sup>1</sup>

### Abstract

Developing a minimally invasive and cost-effective prescreening strategy for colon cancer is critical because of the impossibility of conducting colonoscopy on the entire at-risk population. The concept of field carcinogenesis, in which normal-appearing tissue away from a tumor has molecular and, consequently, nano-architectural abnormalities, offers one attractive approach to identify high-risk patients. In this study, we investigated whether the novel imaging technique partial wave spectroscopic (PWS) microscopy could risk-stratify patients harboring precancerous lesions of the colon, using an optically measured biomarker ( $L_d$ ) obtained from microscopically normal but nanoscopically altered cells. Rectal epithelial cells were examined from 146 patients, including 72 control patients, 14 patients with diminutive adenomas, 20 patients with nondiminutive/nonadvanced adenomas, 15 patients with advanced adenomas/high-grade dysplasia, 12 patients with genetic mutation leading to Lynch syndrome, and 13 patients with cancer. We found that the  $L_d$  obtained from rectal colonocytes was well correlated with colon tumorigenicity in our patient cohort and in an independent validation set of 39 additional patients. Therefore, our findings suggest that PWS-measured  $L_d$  is an accurate marker of field carcinogenesis. This approach provides a potential prescreening strategy for risk stratification before colonoscopy. *Cancer Res*; 72(11); 2720–7. ©2012 AACR.

### Introduction

Colorectal cancer (CRC) is the second leading cause of cancer-related deaths both in the United States and worldwide with a low 5-year survival rate (~60%). There are approximately 141,210 new cases and 49,380 deaths reported in 2011 in United States (1). CRC can be cured if detected at an early stage. However, the early-stage disease is mostly asymptomatic; hence, approximately two thirds of patients with CRCs are diagnosed at a more advanced stage. According to existing guidelines, every individual older than 50 years is a candidate for colonoscopy (2). At present, colonoscopy is considered the gold standard for CRC screening because of its high sensitivity (~97%) to advanced neoplastic lesions and ability to reduce CRC incidence by 65% to 90% by removal of precursor lesions (2). In spite of the unequivocal benefits of colonoscopic exam-

ination, only about one fourth of the eligible screening population undergoes endoscopic CRC screening (3). The reasons for the low compliance rate include the discomfort of the endoscopic procedure, expense, and risk of complications. Even if compliance can be improved, it would be implausible to conduct colonoscopy on the entire at-risk population (≥100 million Americans) given resource constraints. Moreover, it appears that only 20% to 30% of patients harbor neoplasia and only approximately 5% of them are screen-relevant, thus resulting in a majority of colonoscopies being retrospectively unnecessary (4). Therefore, developing a prescreen to colonoscopy is critical. However, currently available screening techniques, such as flexible rectosigmoidoscopy, fecal occult blood test (FOBT), fecal DNA analysis, are either suboptimal in their sensitivity to significant neoplasia or suffer from a high false-positive or false-negative rate (5–7). For example, FOBT has a poor sensitivity (~11%) to detect proximal neoplasia (5), whereas fecal DNA lacks the sensitivity to advanced adenoma (~27%; ref. 7). Similarly, flexible sigmoidoscopy lacks the ability to detect proximal lesions (8), resulting in false-negatives. Most importantly, the inconvenience of bowel preparation can strongly dissuade patients from undergoing alternative techniques such as computed tomographic (CT) colonography (9). Thus, developing a simple, minimally invasive, sufficiently sensitive, and cost-effective precolonoscopic risk stratification technique would be of paramount importance.

One emerging modality of cancer risk stratification is via identification of field carcinogenesis. This represents the impact of the field-of-injury concept—that the genetic and

**Authors' Affiliations:** <sup>1</sup>Biomedical Engineering Department, Northwestern University; <sup>2</sup>Department of Internal Medicine, NorthShore University HealthSystem, Evanston, Illinois; <sup>3</sup>Fox Chase Cancer Center, Philadelphia, Pennsylvania; <sup>4</sup>Division of Gastroenterology and Hepatology, Indiana University School of Medicine, Bloomington, Indiana; and <sup>5</sup>Creighton University Medical Center-Saint Joseph Hospital, Omaha, Nebraska

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

**Corresponding Author:** Hemant K. Roy, NorthShore University HealthSystem, 2650 Ridge Avenue, Suite G208, Evanston, IL 60201. Phone: 847-570-3115; Fax: 847-733-5041; E-mail: h-roy@northwestern.edu

doi: 10.1158/0008-5472.CAN-11-3807

©2012 American Association for Cancer Research.

environmental risk factors confer a fertile mutational field throughout the organ, and the focal neoplastic lesion results from a stochastic mutational event [e.g., truncation of the adenomatous polyposis coli (*APC*) tumor suppressor gene]. This concept is well established in a variety of malignancies (10), such as the diffuse aerodigestive injury associated with smoking-induced lung cancer (11). In the colon, this "condemned" mucosa hypothesis is the rationale for colonoscopic postpolypectomy surveillance (if a patient has had one adenomatous polyp, they are at higher risk of developing others elsewhere in the colon and thus should undergo more frequent colonoscopy). Aside from the adenomatous polyp, there have been a number of putative biomarkers that occur earlier in the predysplastic (i.e., histologically normal) mucosa. These include altered epithelial cell proliferation (12), cell apoptosis (13), gene expression (14), rate of methylation (15), and biochemical (e.g., altered protein kinase C activity; ref. 16). While these all correlate with proximal neoplasia, more accurate markers are needed for clinical implementation.

Our group has been interested in developing nanocytology as a modality of detecting the ultrastructural consequences of the genetic/epigenetic alterations in field carcinogenesis. During field carcinogenesis, the mucosa is normal under light microscopy, but this evaluation is limited to structures above 500 nm due to the diffraction limitation of light. Thus, conventional light microscopy is insensitive to structures, such as ribosomes, macromolecular complexes, higher order chromatin structure, that are in the order of tens to a few hundred nanometers. To probe these nanoscale structures, we have developed a novel optical technology, partial wave spectroscopic microscopy (PWS). PWS is sensitive to structures greater than 10 to 20 nm through analysis of multiple interferences of light reflected from intracellular spatial variations in refractive index and, in principle, is sensitive to essentially any length scale of these variations (limited by signal-to-noise). Because refractive index is a linear function of local macromolecular mass-density (DNA, RNA, proteins), PWS readout is an image of a cell showing the intracellular distribution of a parameter called disorder strength ( $L_d$ ), which quantifies spatial fluctuations in macromolecular density.  $L_d$  is defined as  $L_d = \delta n^\alpha \times l_c^\beta$ , where  $\delta n$  is the standard deviation of the refractive index (and thus mass-density) variations and  $l_c$  is the correlation length of these variations. The coefficient  $\alpha$  depends on the cytology sample preparation ( $\alpha = 1$  in our case), whereas  $\beta$  depends on the configuration of the optical set up and is approximately 1 for the instrumentation used in this study.  $L_d$  is a measure of the spatial variations of macromolecular density and increases with macromolecular condensation (17). The exact nature of the compaction depends on the intracellular location, where  $L_d$  is increased. For example, if  $L_d$  is increased at a particular location in the nucleus, this corresponds to chromatin condensation at that specific location. We have previously reported that  $L_d$  is exquisitely sensitive to subtle genetic/epigenetic perturbations in colon carcinogenesis using colon cancer cell lines and animal models (18). Specifically, microscopically identical but genetically altered (via partial knockdown of the proto-oncogenes or

tumor suppressor genes) CRC cells had aggressiveness parallel with  $L_d$  (18). In animal models of intestinal neoplasia,  $L_d$  was elevated at an early stage (18). Indeed,  $L_d$  increase appears to be a hallmark of field carcinogenesis in colon, lung, and pancreatic cancer (19–21).

Herein, we investigate the clinical potential of PWS interrogation of rectal colonocytes to detect colon field carcinogenesis and hence serve as a minimally intrusive screening technique for colorectal malignancies. We report the study conducted on 146 patients including normal controls, patients harboring adenomatous polyps in their colon, patients having specific genetic mutations leading to Lynch syndrome, and patients with cancer. Our results show a gradient increase in the nano-architectural biomarker, disorder strength ( $L_d$ ), from control to patients having advanced adenoma to patients with cancer. The  $L_d$  increase parallels the patient's risk of developing CRC with respect to different premalignant stages. We also studied the effect of several demographic factors on  $L_d$ . Furthermore, using rectal  $L_d$  as a single marker, we developed the prediction rule in a training set containing controls and patients with advanced adenoma and evaluated its performance on an independent validation set.

## Materials and Methods

### Clinical sample preparation

All studies were conducted and samples were collected with the approval of the Institutional Review Board at NorthShore University HealthSystem (Evanston, IL), Fox Chase Cancer Center (Philadelphia, PA), and Indiana University Medical Center (Bloomington, IN). Patients undergoing screening or surveillance colonoscopy were included in the study. The exclusion criteria included incomplete colonoscopy (failure to visualize cecum), poor colonic preparation, coagulopathy, prior history of pelvic radiation, or systemic chemotherapy. The samples were collected as follows: colonoscopy to cecum was done with standard techniques using Olympus 160 or 180 series or Fujinon colonoscopes. Upon insertion of the colonoscope into the rectum, a cytology brush was passed through the endoscope and gently applied to the visually normal rectum. The brush was then smeared onto a sterile glass slide. The slide was then fixed with 95% ethanol. Although the cytology slide contained different types of cells including epithelial and inflammatory cells, red blood cells, we note that all the measurements reported here were taken from epithelial cells (i.e., colonocytes). This was made possible by staining each patient slide with standardized hematoxylin and cytos-tain staining protocol and directly visualizing the cells before taking the PWS measurements. All the measurements were taken by an operator blinded to the diagnosis and all the colonocytes were selected randomly from the regions not hindered by mucus or cell debris.

### PWS system

The detailed explanation and schematic of the PWS instrument used in this study is reported in the work of Damania and colleagues (22) and in the Supplementary Information. For a given specimen, after normalizing each pixel by the

corresponding incident light profile, a 3-dimensional (3D) data cube  $R(\lambda; x, y)$  is generated, the fluctuating part of the reflection coefficient where  $(x, y)$  refers to a specific pixel in the object plane and  $\lambda$  is the wavelength. The spectral fluctuations in wavelength range from 550 to 700 nm are further analyzed by means of 1D mesoscopic light transport theory to obtain  $L_d$ . Thus, a map of disorder strength  $L_d(x, y)$  is obtained from each pixel  $(x, y)$ . Using this 2D map,  $L_d(x, y)$  for each cell, the mean intracellular disorder strength  $L_d^{(c)}$  (the average over  $x$  and  $y$  pixels), is obtained. The average of  $L_d^{(c)}$  for a group of cells (~50–70 cells for each patient) is calculated and defined as the mean disorder strength per patient,  $L_d^{(p)}$ . In the end, total average is calculated over all the patients of a specific group and that is termed as the group mean of the disorder strength,  $L_d^{(g)}$ , whereas its SD is defined as  $\sigma^{(g)}$ . This average disorder strength  $L_d^{(g)}$  and the SE calculated from its SD,  $\sigma^{(g)}$  are depicted in all the bar plots in this report.

### Statistical Methods

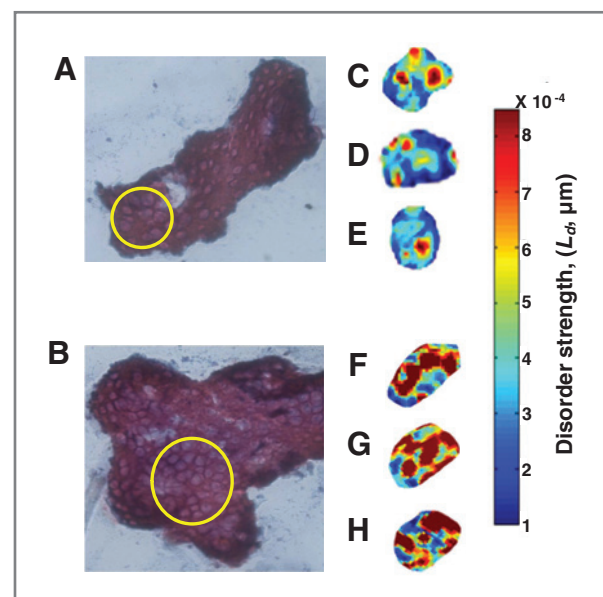
First, we conducted the power analysis to determine the sample size for each high-risk group. For example, patients with advanced adenoma [i.e., high-grade dysplasia (HGD)] are clinically the most screening-relevant population. For this population of patients, we used nQuery Advisor 6.01 Software (Statistical Solutions) and determined that for a sample size of  $n_1 = 72$  (controls) and  $n_2 = 15$  (advanced adenoma), we would have 80% power to detect a between-group difference quantified as an effect size (mean difference divided by common SD) = 0.804, using a 2-sample  $t$  test at a 0.05 significance level. According to Cohen (23), an effect size of 0.80 is considered as a large effect size, which is not unusual to observe in well-controlled experimental studies. In addition, with the sample sizes of 72 and 15 in each of the 2 groups, control and cases, respectively, an area under the ROC (AUROC) curve of 0.80 would be estimated within  $\pm 0.15$  as the 95% confidence interval. This sample size also has more than 88% power to detect the difference between 0.75 and an AUROC curve of 0.50 in null hypothesis, at a 0.05 significance level. Similarly, we conducted sample size analysis for the number of columnar epithelial cells that should be measured for each patient. For the present study, we have found that approximately 40 cells are sufficient to provide a confidence interval on a patient's mean  $L_d$ , that is, 20% of the difference between control and patients with nondiminutive adenoma and 4% for advanced adenomas.

The  $L_d$  obtained in this study had a skewed distribution (a long tail) approximating a log-normal function. Hence, the data were log-transformed to convert to a normal distribution. The Shapiro–Wilk test was used to assess the normality assumption of the log-transformed data, and a  $P$  value  $> 0.05$  provided evidence that the log-transformed data were indeed normally distributed. To accurately measure the performance of  $L_d$ , we calculated the statistical parameters, effect size, and  $P$  values on the log-transformed data. All  $P$  values were calculated with standardized Student  $t$  test on the total number of patients for each subtype. A 2-tailed  $P$  value (assuming unequal variances) of 0.05 or less was considered to be statis-

tically significant in this study. The effect size between 2 groups of patients was calculated on the log-transformed average disorder strength,  $L_d^{(g)}$  and its SD,  $\sigma^{(g)}$ . The value of effect size more than 0.5 is statistically considered significant and it is a more robust parameter than the mean difference between 2 populations. Mean differences provide the percentage fold increase between 2 populations. Moreover, effect size has been used to take into account the slide-to-slide variability and to robustly measure statistical significance of the average disorder strength difference, that is,  $\Delta L_d$  for control group and higher risk patient groups. All  $P$  values and effect sizes were calculated with Microsoft Excel (Microsoft Corporation). Statistical Software Stata (StataCorp LP) was used to generate ANCOVA (analysis of covariance) and AUROC test statistics.

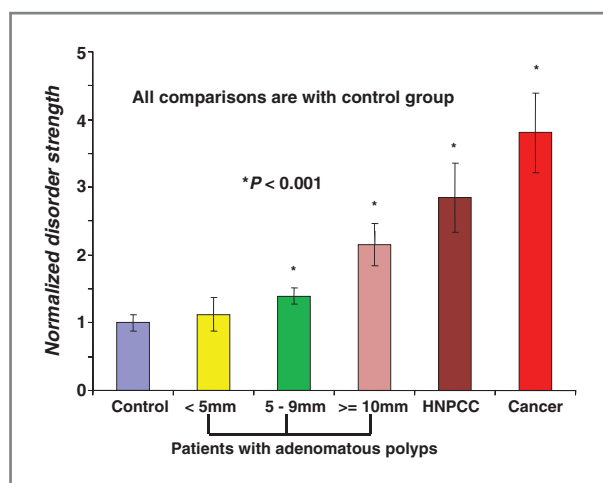
### Results

For each cell, PWS microscopy generates a 2D image of  $L_d(x, y)$ ;  $L_d$  as a function of location within the cells. Figure 1A and B show representative microscope images of stained rectal colonocytes obtained from a control and a patient with cancer, respectively. These images appear microscopically indistinguishable, suggesting no obvious alterations at the microscopic length scales ( $> 300$  nm). However, when the pseudocolor maps of spatial distribution of rectal  $L_d$  are plotted, there appear regions with higher  $L_d$  (represented by red color), indicative of nanoscale perturbations (Fig. 1C–E vs. F–H) for



**Figure 1.** Representative rectal colonocytes from control and cancer patients: A and B are the hematoxylin and eosin-stained microscopic images of rectal epithelium from controls and patients with cancer, respectively. Representative PWS generated pseudocolor heatmap of  $L_d$  for the colonocytes from the circled region of control patient (C–E) and similarly are for patients with cancer (F–H). Although the microscopic images of the colonocytes from the control and patients with cancer are indistinguishable, the disorder strength ( $L_d$ ) was markedly increased in patients with cancer compared with the control.





**Figure 2** Results obtained from  $N = 146$  patients using PWS microscopy show that PWS measured disorder strength ( $L_d$ ) parallels the increasing risk of developing CRCs for patients harboring different size of precancerous lesions (nondiminutive adenoma) in their colon, having specific genetic mutation leading to Lynch syndrome (HNPCC), and frank cancers.

the colonocytes obtained from the patient with cancer compared with control, indicating the nanoscale sensitivity of PWS. Furthermore, the augmentation of rectal  $L_d$  seems to be throughout the cell.

We first investigated whether  $L_d$  was sensitive to field effect in histologically normal-appearing rectal colonocytes obtained from various patient subtypes. In this study, there were  $N = 146$  patients including controls ( $n = 72$ ); patients harboring diminutive adenoma (polyp size  $< 5$  mm,  $n = 14$ ), nondiminutive/nonadvanced adenoma (5–9 mm polyps,  $n = 20$ ), advanced adenoma (polyp size  $\geq 10$  mm, HGD or  $>25\%$  villous features,  $n = 15$ ); patients harboring germ line mutations for Lynch syndrome ( $n = 12$ ), but without concurrent neoplasia; and those having adenocarcinoma in their colon ( $n = 13$ ). Figure 2 and Table 1 show the overall results of rectal  $L_d$  obtained from the reported patient population. As indicated in Fig. 2, there appears to be a progressive increase in  $L_d$  that correlates with the risk of developing CRC: patients with no neoplasia  $<$  nonadvanced adenomas (most of which spontaneously regress)  $<$  advanced adenomas (a more aggressive precancerous lesion, CRC progression risk of 2%–5% per year; ref. 24)  $<$

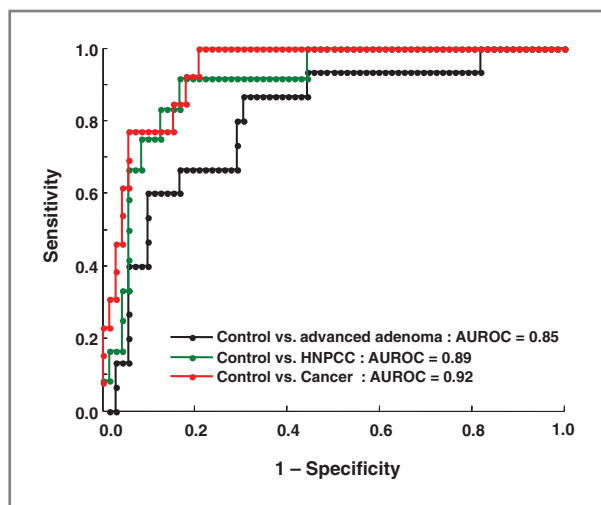
patients with hereditary nonpolyposis CRC (HNPCC; lifetime risk of CRC of 60%–80%; ref. 25)  $<$  patients with frank CRCs. Furthermore, Fig. 2 shows that there is no significant difference in the  $\Delta L_d$  (effect size = 0.08, % difference = 12.44%,  $P \sim 0.68$ ) between control patients and those with diminutive adenoma. This result is consistent with the reported low risk of transformation ( $<0.1\%$ ) of diminutive adenoma to carcinoma (26). However, there was a statistically significant  $\Delta L_d$  (effect size = 0.64, % difference = 42.44%,  $P \sim 0.0001$ ) between control patients and those harboring intermediate size adenomas (5–9 mm polyps).  $\Delta L_d$  increased further (effect size = 1.02, % difference = 113.24%,  $P \sim 0.000006$ ) for patients with advanced adenomas. These results indicate that  $L_d$  is both significantly different and progressively increasing with higher polyp size and hence with higher risk of developing CRCs (26). Importantly, they imply that PWS is sensitive to the colon field carcinogenesis (10, 14, 20) and can quantify alterations in histologically normal appearing rectal colonocytes, irrespective of the actual polyp location.

We further assessed the PWS performance in patients with inherited genetic mutations, which lead to familial CRCs, as it accounts for approximately 15% to 20% of all CRCs (27). Specifically, we investigated HNPCC (i.e., Lynch syndrome) cases, which account for approximately 2% to 3% of all patients with CRCs (27). This disease is mainly caused by germ line mutations in the DNA mismatch repair genes such as *MLH1*, *MSH2*, *MSH6* (27). PWS results from normal-appearing colonocytes showed a 2-fold increase and highly statistically significant  $\Delta L_d$  (effect size = 1.17, % difference = 184.5%,  $P \sim 0.000015$ ) in these patients as depicted in Fig. 2. This sharp increase in  $L_d$  parallels the reported elevated lifetime risk ( $\sim 70\%$ ) of developing CRCs in these patients (27, 28). Moreover,  $\Delta L_d$  is the highest (effect size = 1.42, % difference = 281%,  $P \sim 0.0000006$ ) between the control patients and patients with cancer as shown in Table 1. This is expected as the cells obtained from patients with cancer would have undergone the most nano-architectural alterations of all patient categories. Overall, Fig. 2 and Table 1 indicate that the disorder strength increase parallels the risk of developing CRCs, from control patients to those with neoplastic lesions to patients with a proven history of colon cancer. This increasing trend suggests that rectal  $L_d$  is correlated with the tumorigenicity of colon carcinogenesis and highlights the potential of PWS to risk-stratify patients for CRC based on field-effect,

**Table 1.** The statistical performance of the PWS analysis for all these category of patients compared with the controls

	% Difference	Effect size	P
Control vs. diminutive adenoma	12.44	0.09	0.66
Control vs. nondiminutive/nonadvanced adenoma	42.44	0.64	0.0001
Control vs. advanced adenoma	113.24	1.02	0.000006
Control vs. HNPCC	184.50	1.17	0.000015
Control vs. cancer	281.00	1.44	0.00000066

NOTE: The % mean differences, effect size, and the  $P$  values were calculated on the log-normalized  $L_d$  to obtain a normal distribution.

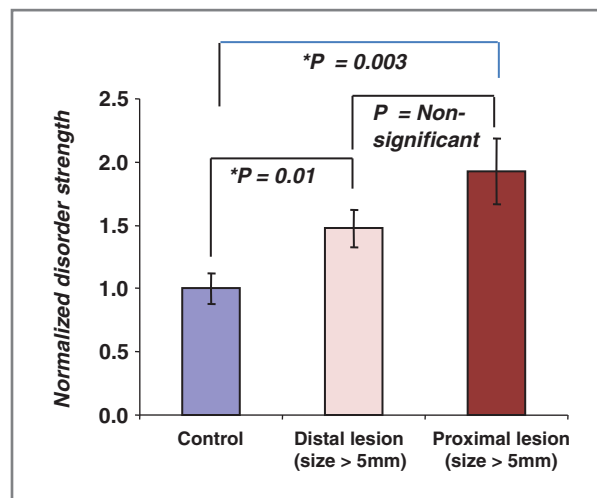


**Figure 3.** The diagnostic performance of the single parameter  $L_d$  for various risk groups. The performance is excellent for the patients with advanced adenoma with AUROC of 0.85, and it improves with AUROC of 0.89 and 0.92 for patients with HNPCCs and frank cancer, respectively.

underscoring it as a promising prescreening technique for colon cancer.

We next evaluated the diagnostic performance of PWS. To gauge the diagnostic capability of the PWS technique, we calculated the performance characteristics using the single parameter  $L_d$ . It is important to note that we used a single marker to avoid any overfitting of the data set presented here. Figure 3 highlights the estimate of the AUROC using  $L_d$  as the only diagnostic marker. The AUROC for PWS analysis of rectal brushings was 0.85 for advanced adenoma/HGD, and it improves further for higher risk patient populations. The AUROC was 0.89 for HNPCCs and 0.92 for patients with cancer. This result highlights the feasibility of the PWS technique for colon cancer screening and the satisfactory power of the single biomarker rectal  $L_d$ . The values of the sensitivity and specificity depicted in Fig. 3 are better (e.g., the sensitivity of FOBT, fecal DNA is ~26% for HGD) than those of other existing screening techniques that are currently (albeit poorly) used as a pre-screen for colonoscopy (6, 7). In the future, these performance characteristics could be further improved by including other independent PWS-derived biomarkers.

One of the challenges with CRC screening is the early detection of proximal colonic neoplasia. There is a growing debate about the sensitivity of colonoscopy to proximal adenomas. Figure 4 provides information about the location of the polyp and its effect on  $L_d$ . It appears from Fig. 4 that there is a statistically nonsignificant difference in  $\Delta L_d$  (effect size = 0.30,  $P \sim 0.15$ ) between patients having proximal and distal lesions. There were 19 patients having proximal polyps of size more than 5 mm and 16 patients with similar distal lesions. These results suggest that PWS is equally sensitive to both proximal and distal lesions. This is an important result with respect to suboptimal efficacy of colonoscopy (29) and flexible sigmoidoscopy (8) to proximal lesions. We believe that PWS



**Figure 4**  $L_d$  from rectal colonocytes is significantly higher compared to controls ( $P$ -value < 0.01) for both, the patients with proximal lesions ( $n = 19$ ) and distal lesions ( $n = 16$ ) of size more than 5 mm. However, there is no significant difference between the proximal and distal lesions themselves ( $P$ -value  $\sim 0.15$ ).

can be a handy tool for clinicians in successfully diagnosing proximal lesions and can potentially reduce the polyp miss rate. In addition, prescreening using PWS could filter patient populations that might benefit from further invasive colonoscopic investigation.

For the control group ( $n = 72$ ) in our study, we included patients harboring hyperplastic polyps ( $n = 7$ ) and patients having diverticulitis ( $n = 8$ ) in their colon. We evaluated the confounding effect of these subtypes of patients on the PWS performance. However, there seems to be no statistically significant difference (effect size = 0.18,  $P \sim 0.25$ ) between  $L_d$  values of the patients with nonneoplastic lesions and those with clean colons. Hence, we combined all these patients in the control group.

We next studied the role of confounding demographic risk factors such as age, gender, smoking, and drinking history on the sensitivity of the measured biomarker. Age has been implicated as one of the key risk factors for colonic neoplasia and there have been a variety of age-related changes in colonic mucosa (such as methylation; ref. 30). We therefore conducted ANCOVA analysis and noted no significant confounding with age ( $P = 0.54$  for  $L_d$ ). As outlined in Table 2, smoking and drinking history also did not have any confounding effect on  $L_d$  ( $P = 0.57$  for current smokers and  $P = 0.99$  for drinking). Similarly, male gender is a well-established risk factor for colonic neoplasia (31). However, ANCOVA analysis indicated that there was no significant confounding with gender ( $P = 0.29$  for  $L_d$ ). Overall, the nonsignificant ANCOVA  $P$  values suggest that  $L_d$  is not confounded by age, gender, smoking, or drinking patterns.

The next question we addressed was the performance of  $L_d$  in a prospective study. Although we understand that we have a modest data set to make any definitive conclusions, we still wanted to gauge the diagnostic power of our approach when tested on an independent training and testing set for any patient having an adenomatous polyp of size 10 mm or more

**Table 2.** The impact of demographic factors on the single biomarker

Demographic factors	Control	Diminutive adenoma	Nondiminutive/nonadvanced adenoma	Advanced adenoma	Lynch syndrome or HNPCC	Cancer	Effect on rectal $L_d$ , ANCOVA $P$
Age (mean $\pm$ SD), y	55 $\pm$ 9	65 $\pm$ 13	58 $\pm$ 10	65 $\pm$ 13	43 $\pm$ 12	66 $\pm$ 13	0.54
Gender (% male)	49	64	73	42	42	36	0.20
Smoking (% smokers)	12	0	33	8	25	14	0.57
Drinking (% alcoholic)	70	64	73	58	67	57	0.99

NOTE: Demographic characteristics such as age, gender, smoking, and drinking history is shown for different patient groups and their effect on  $L_d$ . The  $P$  value is calculated using the ANCOVA in Stata. It is evident that  $L_d$  is not confounded by age ( $P = 0.54$ ), gender ( $P = 0.20$ ), smoking ( $P = 0.57$ ), and drinking history ( $P = 0.99$ ).

or with tubulovillous features. We specifically selected patients with advanced adenoma, as they comprise the most clinically relevant screening population. The training set was composed of a subset of  $n = 87$  patients (wherein 72 were controls and 15 were patients with advanced adenoma) of the  $N = 146$  total patients. However, the validation set ( $N = 39$  patients with 14 control and 25 with advanced adenoma) was an independent enriched data set for case-control recruited from a second clinical site. We developed a cutoff on the basis of the  $L_d$  values of the training set and applied the same to the validation set. The estimates of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) based on this threshold are listed in Table 3. As summarized in Table 3, we obtained promising performance characteristics for the training set (sensitivity = 73%, specificity = 78%, PPV = 41%, NPV = 93%). The 95% confidence intervals for all these parameters were: sensitivity = 0.733 (0.48–0.891); specificity = 0.778 (0.669–0.858); PPV = 0.407 (0.245–0.593); and NPV = 0.933 (0.841 to 0.974). We then applied the same  $L_d$  threshold to the validation set, which yielded sensitivity = 74% and spec-

ificity = 83%, similar to the training set and further supporting the robustness of our conclusions. We again emphasize that this performance is based on a single biomarker ( $L_d$ ) and it is probable that inclusion of additional PWS-measured markers would further improve the performance characteristics.

### Discussion

Herein, we show that field carcinogenesis-based PWS analysis of rectal colonocytes has promise as a novel, minimally intrusive CRC risk stratification technique. Our results obtained from 146 patients showed rectal  $L_d$  paralleled the risk of developing cancer: diminutive adenoma < nondiminutive/nonadvanced adenoma < advanced adenoma. Intriguingly, the rectal  $L_d$  of patients with HNPCCs without concurrent neoplasia was higher than patients with advanced adenomas, but lower than CRCs, further supporting the notion that  $L_d$  mirrored overall risk of CRC development. The diagnostic promise of this single biomarker was validated with an independent, albeit small, data set. Rectal  $L_d$  appeared robust and not confounded by cosegregating risk factors (age, gender, smoking history, and alcohol consumptions) and was able to sense proximal and distal neoplasia equivalently.

While colonoscopy is the recommended screening option for CRCs (2), it cannot be applied on the entire at-risk population due to cost, resource constraints, and possible complications. Moreover, some of the current noninvasive screening options have poor sensitivity to neoplasia. For example, guaiac-based FOBT has sensitivity of 10.8% (6), whereas fecal DNA has sensitivity of 27% (5, 7) to clinically significant lesions. In this situation, field cancerization provides a possible solution for developing a minimally invasive prescreen. There are several reports supporting the biological plausibility of using rectal mucosa to detect CRCs (10). Several lines of evidence suggest that there are early genetic/epigenetic and consequently morphological changes that occur in the rectal mucosa before the development of adenocarcinoma (14, 32, 33). There have been a few efforts to detect CRC using field effect, for example, flexible sigmoidoscopy to detect a sentinel distal adenoma as a marker of advanced proximal neoplasia, however its sensitivity is low (~33% in women; ref 29). In contrast, the PWS nanocytology-based approach is minimally invasive, easy, quick, cost-effective, more patient-compliant, and

**Table 3.** The performance of  $L_d$  on an independent training ( $n = 87$ ) and testing set ( $n = 39$ ) developed for controls and patients having advanced adenoma

	Training set (%); N = 87 patients	Validation set (%); N = 39 patients
Sensitivity	73	74
Specificity	78	83
PPV	41	N/A
NPV	93	N/A

NOTE: Sensitivity, specificity, PPV, NPV are shown for controls and patients having advanced adenoma in their colon using the training set. On the basis of a single cutoff  $L_d$  value derived from the training set, we obtained equivalent performance characteristics (i.e., sensitivity and specificity) for the enriched testing set. As the testing data set is an enriched population, predictive values (both negative and positive) are uninformative.

sufficiently sensitive (AUROC of 0.85 for patients with advanced adenoma, 0.89 for HNPCC, and 0.92 for cancer) for clinical practice. Further studies using optimized instrumentation and potentially more PWS markers will likely show improved diagnostics. Hence, PWS nanocytology has the promise of translating a field carcinogenesis approach into clinically practical means of colonic risk stratification.

Recently, CT colonography (virtual colonoscopy) has been sanctioned for average risk screening (9). In comparison, the performance of rectal PWS data had similar per-polyp rate sensitivity (~90%) for advanced adenomas to the reported large multicenter trials using CT colonography. However, PWS nanocytology has advantages, including a lack of need for colon purge, less discomfort and expense, and no radiation exposure. This is likely to improve patient compliance with screening, which has been a major barrier to reducing the toll of CRCs.

Some recent reports discuss the possibility of functional proteomics for diagnosing CRCs in clinical practice (34). They describe alterations in the expression of 9 proteins in precancerous and neoplastic tissues, for example, nm23, manganese superoxide dismutase (MnSOD), suggesting their role in colon tumorigenesis (16, 35). Although promising, the proteomic approaches lack reproducibility, suffer from tedious sample analysis (using mass spectrometry) and conflicting results for different biomarkers (34). Hence, their translation into clinical practice is a big question mark (34). However, in comparison, our PWS-driven screening approach (which investigates changes at similar length scales) provides compelling results with satisfactory sensitivity and specificity for patients with advanced histologies in the colon.

While the rectal  $L_d$  appears to be a more accurate and robust biomarker than conventional approaches, the biological underpinnings have been incompletely elucidated. Others have shown that there are subtle submicron abnormalities in both localized and more distal field carcinogenesis (33, 36, 37). With regard to PWS,  $L_d$  represents changes in the local mass density of cellular building blocks (proteins, RNA, DNA) with an increase in  $L_d$  corresponding to macromolecular condensation (e.g., condensation leads to higher local mass density and correlation length, hence higher  $L_d$ ). These changes occur both in the cytoplasm and nucleus. Our studies, investigating the cytoplasmic origin, have taken a candidate approach focusing on the cytoskeleton because of early proteomic data (38, 39) and the observation that many of the key molecules in early colon carcinogenesis actually interact with the cytoskeleton (APC,  $\beta$ -catenin, E-cadherin, Src). We have recently reported that pharmacologic disruption of the cytoskeleton ameliorated the proneoplastic (increase)  $L_d$  in stably transfected colon cancer cell lines (22). From a nuclear perspective,  $L_d$  increase indicates chromatin condensation, which was further corroborated by our group's electron microscopy demonstration that histologically normal rectal biopsies from patients harboring adenomas had altered chromatin architecture (e.g., higher heterochromatin content) and exhibited chromatin clumping, potentially driven by histone deacetylase 2 (HDAC2) overexpression (17, 40, 41). This finding is consonant with the recent interest in high-order chromatin structure as a surrogate for cellular transcriptional activity.

There are several limitations with our approach that need to be acknowledged. First, we report an outstanding performance of rectal PWS marker,  $L_d$  that is tested only on a modest sample size. We realize that to validate our approach, there is a need to conduct studies over a large data set of the prospective population. Second, there are other rectal diseases (e.g., inflammatory bowel disease, ulcerative colitis etc.), which can increase the risk of CRCs (42) but are not a part of this study. Future studies will include this group of patients. Third, we report performance based on a single marker, rectal  $L_d$ ; however, identifying more PWS-derived markers can improve the performance characteristics.

In conclusion, we provide evidence that interrogation of the rectal epithelium with PWS can quantify nanoscale architectural alterations in colon field carcinogenesis. The current approach appears to have reasonable diagnostic accuracy (comparable with standard techniques) and will likely improve with ongoing technological refinements. If confirmed in large-scale validation trials, we envision that it would be applied as a minimally invasive and cost-effective prescreen technique that can identify patients at high-risk and who would likely benefit from further testing (colonoscopy). Using PWS nanocytology with field carcinogenesis for accurate risk stratification could be a paradigm shift in CRC screening.

#### Disclosure of Potential Conflicts of Interest

H.K. Roy has employment (other than primary affiliation; e.g., consulting) in Small Business as the owner, ownership interest (including patents) in Small business/American Bio Optics LLC and Nanocytomics, and has other (e.g., expert testimony) in Small Business. H. Subramanian has employment (other than primary affiliation; e.g., consulting) in Nanocytomics for consulting. M.J. Goldberg has employment (other than primary affiliation; e.g., consulting) in abo as founder and ownership interest (including patents) in abo. V. Backman has ownership interest (including patents) in Nanocytomics, LLC. No potential conflicts of interests were disclosed by the other authors.

#### Authors' Contributions

**Conception and design:** D. Damanian, H.K. Roy, H. Subramanian, D.S. Weinberg, M.J. Goldberg, J. Muldoon, V. Backman

**Development of methodology:** H. Subramanian, M.J. Goldberg, J. Muldoon, L. Cherkezyan, Y. Zhu, P. Pradhan, V. Backman

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** D. Damanian, H.K. Roy, D.S. Weinberg, D.K. Rex, M.J. Goldberg, J. Muldoon, Y. Zhu, L.K. Bianchi, D. Shah, M. Borkar, H. Lynch

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** D. Damanian, H.K. Roy, D.S. Weinberg, L. Cherkezyan

**Writing, review, and/or revision of the manuscript:** D. Damanian, H.K. Roy, H. Subramanian, D.S. Weinberg, D.K. Rex, P. Pradhan, V. Backman

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** H. Lynch, V. Backman

**Study supervision:** D. Damanian, H.K. Roy, H. Subramanian, V. Backman

#### Acknowledgments

The authors thank Andrew Gomes and Hongyan Du for their valuable suggestions during statistical discussion.

#### Grant Support

This research work was supported by the NIH grants R01CA128641, R01EB003682, U01 CA111257, and National Science Foundation grant CBET-0937987.

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

Received November 29, 2011; revised March 7, 2012; accepted March 26, 2012; published OnlineFirst April 6, 2012.



## References

- Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin* 2011;61:212–36.
- Levin B, Lieberman DA, McFarland B, Andrews KS, Brooks D, Bond J, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology* 2008;134:1570–95.
- Cooper GS, Doug Kou T. Underuse of colorectal cancer screening in a cohort of Medicare beneficiaries. *Cancer* 2008;112:293–9.
- Roy HK, Backman V, Goldberg MJ. Colon cancer screening: the good, the bad, and the ugly. *Arch Intern Med* 2006;166:2177–9.
- Morikawa T, Kato J, Yamaji Y, Wada R, Mitsushima T, Shiratori Y. A comparison of the immunochemical fecal occult blood test and total colonoscopy in the asymptomatic population. *Gastroenterology* 2005;129:422–8.
- Imperiale TF, Ransohoff DF, Itzkowitz SH, Turnbull BA, Ross ME. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. *N Engl J Med* 2004;351:2704–14.
- Song K, Fendrick AM, Ladabaum U. Fecal DNA testing compared with conventional colorectal cancer screening methods: a decision analysis. *Gastroenterology* 2004;126:1270–9.
- McCallion K, Mitchell RM, Wilson RH, Kee F, Watson RG, Collins JS, et al. Flexible sigmoidoscopy and the changing distribution of colorectal cancer: implications for screening. *Gut* 2001;48:522–5.
- Johnson CD, Chen MH, Toledano AY, Heiken JP, Dachman A, Kuo MD, et al. Accuracy of CT colonography for detection of large adenomas and cancers. *N Engl J Med* 2008;359:1207–17.
- Kopelovich L, Henson DE, Gazdar AF, Dunn B, Srivastava S, Kelloff GJ, et al. Surrogate anatomic/functional sites for evaluating cancer risk: an extension of the field effect. *Clin Cancer Res* 1999;5:3899–905.
- Bhutani M, Pathak AK, Fan Y-H, Liu DD, Lee JJ, Tang H, et al. Oral epithelium as a surrogate tissue for assessing smoking-induced molecular alterations in the lungs. *Cancer Prev Res (Phila)* 2008;1:39–44.
- Anti M, Marra G, Armelao F, Percesepe A, Ficarelli R, Ricciotti GM, et al. Rectal epithelial cell proliferation patterns as predictors of adenomatous colorectal polyp recurrence. *Gut* 1993;34:525–30.
- Bernstein C, Bernstein H, Garewal H, Dinning P, Jabi R, Sampliner RE, et al. A bile acid-induced apoptosis assay for colon cancer risk and associated quality control studies. *Cancer Res* 1999;59:2353–7.
- Chen LC, Hao CY, Chiu YS, Wong P, Melnick JS, Brotman M, et al. Alteration of gene expression in normal-appearing colon mucosa of APC(min) mice and human cancer patients. *Cancer Res* 2004;64:3694–700.
- Ushijima T. Epigenetic field for cancerization. *J Biochem Mol Biol* 2007;40:142–50.
- Polley AC, Mulholland F, Pin C, Williams EA, Bradburn DM, Mills SJ, et al. Proteomic analysis reveals field-wide changes in protein expression in the morphologically normal mucosa of patients with colorectal neoplasia. *Cancer Res* 2006;66:6553–62.
- Kim JS, Pradhan P, Backman V, Szleifer I. The influence of chromosome density variations on the increase in nuclear disorder strength in carcinogenesis. *Phys Biol*. 2011;8:015004.
- Subramanian H, Pradhan P, Liu Y, Capoglu IR, Li X, Rogers JD, et al. Optical methodology for detecting histologically unapparent nanoscale consequences of genetic alterations in biological cells. *Proc Natl Acad Sci U S A* 2008;105:20118–23.
- Subramanian H, Pradhan P, Liu Y, Capoglu IR, Rogers JD, Roy HK, et al. Partial-wave microscopic spectroscopy detects subwavelength refractive index fluctuations: an application to cancer diagnosis. *Opt Lett* 2009;34:518–20.
- Subramanian H, Roy HK, Pradhan P, Goldberg MJ, Muldoon J, Brand RE, et al. Nanoscale cellular changes in field carcinogenesis detected by partial wave spectroscopy. *Cancer Res* 2009;69:5357–63.
- Roy HK, Subramanian H, Damania D, Hensing TA, Rom WN, Pass HI, et al. Optical detection of buccal epithelial nanoarchitectural alterations in patients harboring lung cancer: implications for screening. *Cancer Res* 2010;70:7748–54.
- Damania D, Subramanian H, Tiwari AK, Stypula Y, Kunte D, Pradhan P, et al. Role of cytoskeleton in controlling the disorder strength of cellular nanoscale architecture. *Biophys J* 2010;99:989–96.
- Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.; 1988.
- Brenner H, Hoffmeister M, Stegmaier C, Brenner G, Altenhofen L, Haug U. Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. *Gut* 2007;56:1585–9.
- Lynch HT, de la Chapelle A. Hereditary colorectal cancer. *N Engl J Med* 2003;348:919–32.
- Butterly LF, Chase MP, Pohl H, Fiarman GS. Prevalence of clinically important histology in small adenomas. *Clin Gastroenterol Hepatol* 2006;4:343–8.
- Lynch HT, de la Chapelle A. Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet* 1999;36:801–18.
- Brand RM, Jones DD, Lynch HT, Brand RE, Watson P, Ashwathnayan R, et al. Risk of colon cancer in hereditary non-polyposis colorectal cancer patients as predicted by fuzzy modeling: influence of smoking. *World J Gastroenterol* 2006;12:4485–91.
- Lewis JD, Ng K, Hung KE, Bilker WB, Berlin JA, Brensinger C, et al. Detection of proximal adenomatous polyps with screening sigmoidoscopy: a systematic review and meta-analysis of screening colonoscopy. *Arch Intern Med* 2003;163:413–20.
- Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JP. Aging and DNA methylation in colorectal mucosa and cancer. *Cancer Res* 1998;58:5489–94.
- Roy HK, Gomes AJ, Ruderman S, Bianchi LK, Goldberg MJ, Stoyneva V, et al. Optical measurement of rectal microvasculature as an adjunct to flexible sigmoidoscopy: gender-specific implications. *Cancer Prev Res (Phila)* 2010;3:844–51.
- Kukitsu T, Takayama T, Miyanishi K, Nobuoka A, Katsuki S, Sato Y, et al. Aberrant crypt foci as precursors of the dysplasia-carcinoma sequence in patients with ulcerative colitis. *Clin Cancer Res* 2008;14:48–54.
- Alberts DS, Einspahr JG, Krouse RS, Prasad A, Ranger-Moore J, Hamilton P, et al. Karyometry of the colonic mucosa. *Cancer Epidemiol Biomarkers Prev* 2007;16:2704–16.
- de Noo ME, Tollenaar RA, Deelder AM, Bouwman LH. Current status and prospects of clinical proteomics studies on detection of colorectal cancer: hopes and fears. *World J Gastroenterol* 2006;12:6594–601.
- Reymond MA, Steinert R, Kahne T, Sagynaliev E, Allal AS, Lippert H. Expression and functional proteomics studies in colorectal cancer. *Pathol Res Pract* 2004;200:119–27.
- Weyn B, Jacob W, da Silva VD, Montironi R, Hamilton PW, Thompson D, et al. Data representation and reduction for chromatin texture in nuclei from premalignant prostatic, esophageal, and colonic lesions. *Cytometry* 2000;41:133–8.
- Verhest A, Kiss R, Dolne D, Larsimont D, Salmon I, Delaunoy Y, et al. Characterization of human colorectal mucosa, polyps, and cancers by means of computerized morphonuclear image analyses. *Cancer* 1990;65:2047–54.
- Bernstein C, Bernstein H, Payne CM, Dvorak K, Garewal H. Field defects in progression to gastrointestinal tract cancers. *Cancer Lett* 2008;260:1–10.
- Bernstein H, Payne CM, Kunke K, Crowley-Weber CL, Waltmire CN, Dvorakova K, et al. A proteomic study of resistance to deoxycholate-induced apoptosis. *Carcinogenesis* 2004;25:681–92.
- Pradhan P, Damania D, Joshi HM, Turzhitsky V, Subramanian H, Roy HK, et al. Quantification of nanoscale density fluctuations using electron microscopy: light-localization properties of biological cells. *Appl Phys Lett* 2010;97:243704.
- Stypula Y, Crawford S, Subramanian H, Tiwari AK, Kunte D, Cornwell ML, et al. Understanding biological mechanisms of nuclear disorder strength in early carcinogenesis. *Gastroenterology* 2011;140:S765–6.
- Sands BE, Grabert S. Epidemiology of inflammatory bowel disease and overview of pathogenesis. *Med Health R I* 2009;92:73–7.