Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Detecting and Quantifying Biomarkers of Risk for Colorectal Cancer Using Quantum Dots and Novel Image Analysis Algorithms

Roberd M. Bostick[1], Koon Yin Kong, Thomas U. Ahearn, Qaiser Chaudry, Vaunita Cohen, and May D. Wang

***Abstract*—Colorectal cancer, the second leading cause of cancer deaths in the United States, is a molecular disease that is largely lifestyle determined and preventable. While heart disease has been sharply declining, in large part from widespread use of biological measurements that indicate risk ("biomarkers of risk"), such as blood cholesterol, to motivate and guide preventive treatment, colorectal cancer is a disease for which mortality rates have changed little and for which there have been no biomarkers of risk. Based on new knowledge about the molecular basis of colorectal cancer we developed and validated a panel of treatable biomarkers of risk that can be measured in rectal biopsies using automated immunohistochemistry and semi-automated image analysis. The methodology is now being made practical for clinical application through the use of 1) quantum dots, so that all of the biomarkers can be detected simultaneously on the same histologic sections (ie, multiplexed), and 2) novel, automated image analysis algorithms to measure the quantities and tissue distributions of the biomarkers. Herein we summarize our methods, results, current directions, and progress.**

I. Introduction

Cancer has become the leading cause of death in persons 85 years of age or less in the US [1]. Colorectal cancer is the second leading cause of cancer deaths after lung cancer, and is the only major cancer to affect men and women essentially equally [1][2]. International rates for colorectal cancer vary 20-fold [2], and migrants from low risk countries to the US (which has the highest rates) develop the rates found in the US within a generation, attesting to the profound role of lifestyle in the etiology of the disease, and thus to its preventability [2]. However, mortality from colorectal cancer has changed very little over the past 50 years [1][2]. Meanwhile, mortality from cardiovascular disease, the former number one cause of death in the US, has dramatically declined over the past 30 years [1][3], primarily because biomarkers of risk (biological measurements that indicate risk for a future disease) were developed and validated, such as blood lipid profiles (total, LDL-, and HDL-cholesterol, etc) and blood pressures, which could be treated to prevent future heart disease [3]. There are currently no such biomarkers of risk for colorectal or any other cancer. Whereas biomarkers of risk for ischemic heart disease are based on measuring key elements of the known pathophysiology of the disease, an analogous approach was not possible until recently for colon cancer. Over the past ten years the molecular basis for the disease has become quite clear [4], [5], although much remains to be learned. This knowledge has not been translated into methods for cancer risk prediction and prevention.

II. Previous Work

Based on what is now known of the molecular basis of colon cancer, we selected 19 key elements to measure as potential molecular phenotypic biomarkers. We successfully developed automated immunohistochemical (IHC) methods to detect the biomarkers in normal colon tissue (Fig. 1), and

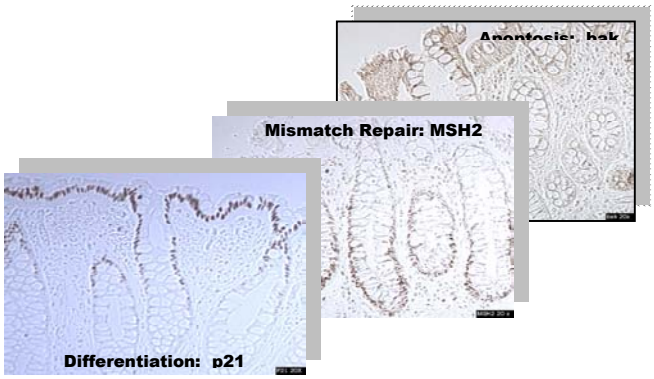

Fig. 1. Examples of biomarkers of risk for colorectal cancer detected immunohistochemically in the colon crypts of normal-appearing colorectal mucosa obtained by rectal biopsy.

a custom image analysis software program to measure the levels of the biomarkers in the normal colon tissue. We also conducted a community-, colonoscopy-based case-control study of incident, sporadic colorectal adenoma (the immediate precursor to the vast majority of colon cancers) to compare biomarker levels in persons at low risk (no colon tumors now or previously) vs. high risk (current, but no previous, colon tumor). Measurement reliability was high (test-retest reliability correlation coefficients: 0.94 – 0.99). Examples of the stronger findings related to validity included that in cases APC was 25% lower, the APC/E-cadherin ratio was 40% lower, TGFα was 37% higher,

[1] Manuscript received June 19, 2006. This work was supported in part by grants from the National Institutes of Health (CA104637 and CA119338 to R. M B., and GM072069, CA108468, and CA119338 to M. D. W.) and Georgia Cancer Coalition Distinguished Scholar awards to R. M. B. and M. D. W.

R. M. Bostick, Department of Epidemiology and Winship Cancer Institute, Emory University, Atlanta, GA 30322 USA (phone: 404-727-2671; fax: 404-727-8737; email: rmbosti@sph.emory.edu).

M. D. Wang, Electrical and Computer Engineering, The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA 30332 USA (email: maywang@bme.gatech.edu).

K. Y. Kong and Q. Chaudry, Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (kykong@gatech.edu, qaiser@gatech.edu).

T. U. Ahearn and V. Cohen, Department of Epidemiology, Emory University, Atlanta, GA 30322 USA (tahearn@sph.emory.edu, vcohen@sph.emory.edu).

$TGF\beta_1$ was 33% lower, and the bax/bcl-2 ratio was 26% higher. Differences of 25 – 40% applied to total cholesterols translates to total cholesterol values of approximately 250 to 280 mg/100 ml, values that substantially increase risk for ischemic heart disease and that are vigorously treated with diet and exercise and/or medication to lower risk [3]. Additional analyses revealed that our biomarkers were strongly associated with several lifestyle factors, suggesting that the biomarkers may be treatable using these lifestyle factors as interventions.

Although our findings are important and useful in the research setting, our current laboratory methods are impractical for clinical applications. First, we use IHC methods to detect the biomarkers in histologic slides of normal colon tissue (obtained in a procedure somewhat comparable to a Pap smear – no preparation is required of the patient and a superficial sampling of colon epithelium is obtained from the rectum about the same distance in as a digital rectal exam in a painless, two-minute outpatient procedure). To detect six different biomarkers, a set of five slides must be processed for each biomarker separately, for a total of 30 slides (6 biomarkers x 5 slides per set); each set of slides must also be analyzed separately. With our new nanotechnology-based methods (quantum dot-based IHC, described in more detail further below), all six biomarkers can be detected at the same time ("multiplexed") on one set of five slides (Fig. 2). Second, we developed and use an

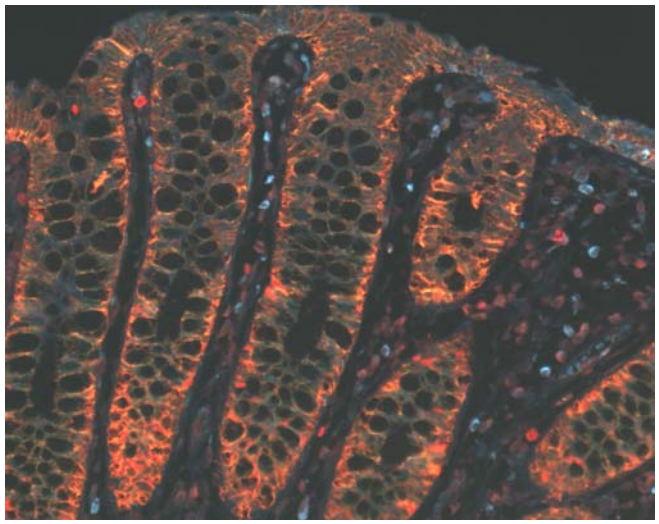

Fig. 2. Example of three biomarkers of risk for colorectal cancer detected with quantum dot immunohistochemistry in colon crypt in normal-appearing colorectal mucosa obtained by rectal biopsy.

image analysis program to quantify the detected expression of the biomarkers on the slides. This program is the first to allow measurement of the total quantity of expression and the tissue distribution of the biomarkers in normal colon tissue. Although the program makes the measurements of the biomarkers possible, it still requires a substantial amount of operator input, and since it also must interface with multiple other software programs, it requires a substantial maintenance effort. Currently, it takes about seven hours to analyze six biomarkers. The program is currently being modified (described in more detail further below) so that it can be used on slides processed by quantum dot immunohistochemistry (Q-Dot IHC), analyze six or more biomarkers on a slide at the same time, incorporate a new pattern recognition algorithm that will eliminate most operator input involved with slide evaluation, and be a standalone program. We estimate that the efficiencies gained by combining our new nanotechnology-based detection and image analysis methods will decrease slide analysis time to a total of 20 minutes, making them suitable for clinical use.

III. Biomarker Detection with Quantum Dots

In outline, immunohistochemistry (IHC) is a procedure in which an antigen (e.g., a protein biomarker) in a tissue is identified in a series of steps (Fig. 3), including the application of a primary antibody to the antigen, linking the primary antibody to a secondary antibody that has an attachment site for a chemical linking agent, adding the linking agent that has an attachment site for a chromogen (e.g., DAB), and then adding the chromogen (one can also apply a counterstain at this point). Only one antigen can be detected on a given slide using this method. Our new method uses specially coated nanocrystals, called quantum dots (Q-Dots) , instead of the chromogen (Fig. 3).

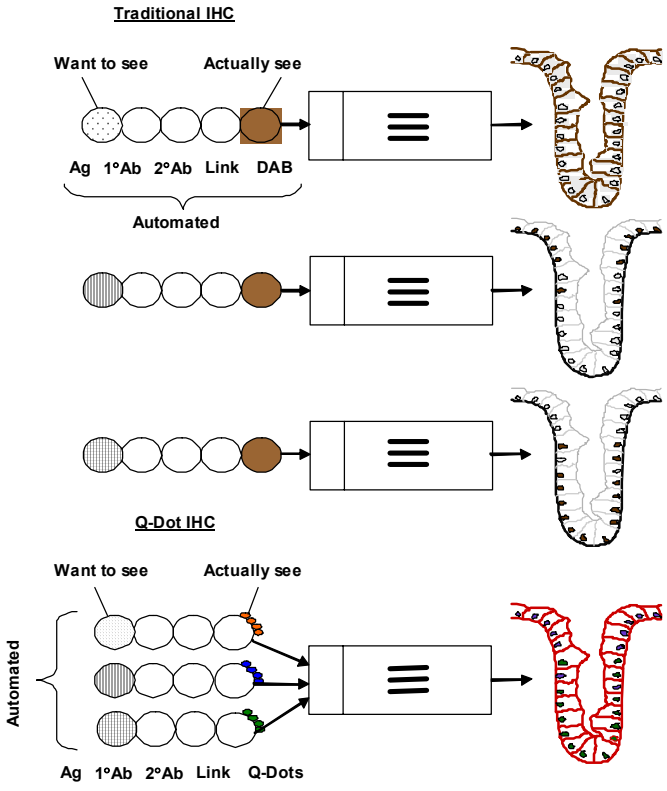

Fig. 3. Schematic representation of traditional and quantum dot immunohistochemistry to detect biomarkers of risk for colorectal cancer in normal colon crypts.

Q-Dots have the property of being excited by any type or wavelength of light to emit light in a very narrow spectrum.

Q-Dots of slightly different sizes emit different, non-overlapping spectra. Q-Dots can be conjugated to the usual linking agents used in traditional IHC (Fig. 3). This means that we can link Q-Dot-linking agent complexes with different size Q-Dots to different antibodies and thus to different antigens, thereby allowing detection of multiple biomarkers on the same slides (we term this "Q-Dot-IHC") (Fig. 3). Also, in contrast to immunofluorescent dyes, the light emissions from quantum dots last months rather than just a few minutes, thus making analysis feasible in population- or clinical-based studies. We have found that no counterstain is needed and that six different biomarkers can be detected simultaneously with our current methods. Although we have proved that we can identify biomarkers singly and multiplexed using our new methods, there are numerous variables involved in our Q-Dot-IHC protocols; consequently, a substantial effort is underway to optimize the protocols, especially the final multiplexed protocol. This effort to finalize the multiplexed protocol combining the six biomarkers involves varying the antibody dilutions and application times, varying the Q-Dot concentrations, and linking the different Q-Dot-linking agent complexes to different antibodies, etc. Through multiple experiments we are identifying the sources of experimental error and ways to control them (e.g., effects of storage conditions, such as time, light exposure, temperature, etc.).

IV. Quantum Dot-Labeled Biomarker Image Analysis and Quantification

The software program that we developed for quantifying biomarkers in normal colon tissue on immunohistochemically processed histologic slides was a major step forward in colon cancer biomarker assessment. Formerly, biomarkers could only be assessed subjectively or at best semi-quantitatively. All of our colon cancer risk biomarkers are expressed in density gradients along colon crypts (key structures of interest in the tissue lining the colon) and thus cannot be quantified visually. Our image analysis program overcomes two major hurdles. First, since the amount of staining of a biomarker is proportional to the amount of biomarker in the tissue, and since the optical density of the staining is proportional to the amount of staining, we can quantify the amount of biomarker in the tissue using optical density measurements. Second, our program incorporates a drawing tool that allows us to outline the colon crypts in digital images of the colon tissue, and then the program, maintaining spatial orientation, automatically divides each crypt into equal sized segments from the colon crypt top to bottom, and then measures the optical density of the staining in each segment, thus providing data to quantify the distribution of the biomarker in the colon crypts (Fig. 4).

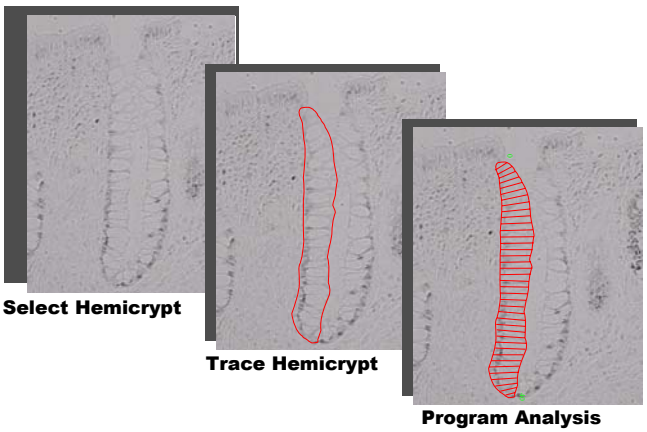

Fig. 4. Example of application of algorithm in image analysis program to quantify biomarkers of risk for colorectal cancer detected immunohistochemically in the colon crypts of normal-appearing colorectal mucosa obtained by rectal biopsy.

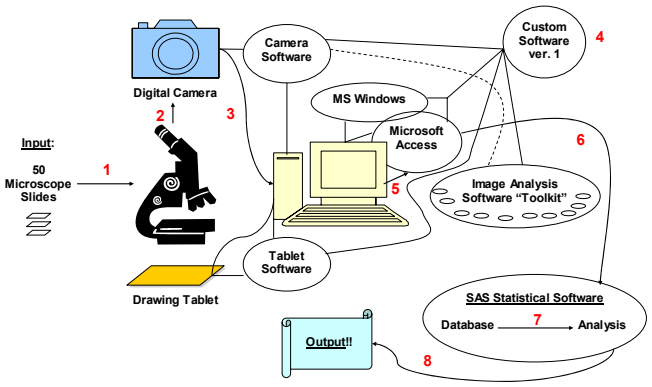

Fig. 5. Current image analysis system for slides processed by traditional immunohistochemistry.

The system to accomplish this is depicted in Fig. 5. Five slides per biomarker must be viewed under a light microscope one a time, and multiple images on each slide must viewed one at a time. So, for example, if six biomarkers are to be analyzed, a total of 30 slides must be viewed. Each image to be analyzed from each slide must be acquired with a digital camera, and then captured and transferred to the computer hard disc one at a time. Each image from each slide is processed using our custom software which supplies our colon crypt analysis algorithm and coordinates needed functions of MS Windows, the digital camera and digital drawing palette drivers, ImagePro Plus (a commercially available image analysis toolkit software package by Media Cybernetics), and Microsoft Access. Raw image data are deposited in Microsoft Access under the control of our custom software. The raw database must then be transferred from Microsoft Access into a statistical software (such as SAS) dataset where final derived analysis variables are created, and then analyzed and reported out. Despite the automated features, extensive manual involvement is required. The entire process for six biomarkers takes about seven hours. In addition, effort to maintain the system is substantial.

As depicted in Fig. 6, in the advanced image analysis system for Q-Dot IHC that we are developing, to analyze the same six biomarkers, a total of only five slides will need to be scanned into a fluorescence slide scanner/virtual microscope system from which digital images of slides will automatically be stored. Multiple sub-images of each slide image will be processed using advanced custom software (see below), which will also create a database and derived

analysis variables, then analyze the data and report the final output. It is estimated that the entire process will take only about 20 minutes during which time the operating technician can be conducting other work. Furthermore, the system will be usable across platforms and the Web and will be easy to maintain.

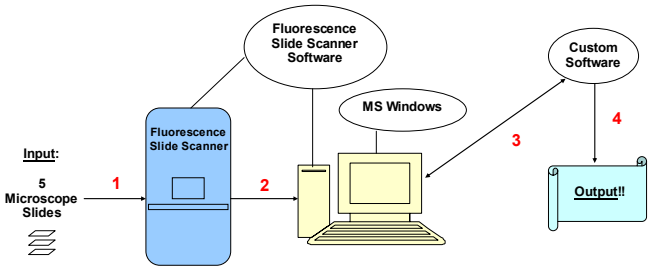

Fig. 6. Advanced image analysis system for quantum dot immunohistochemistry under development.

Now that we have proved that our methodology is valid, for the advanced system outlined above, we are developing and testing novel algorithms to identify and analyze individual cells within the crypts and to eliminate the requirement for manual identification and outlining of crypts. As examples, we developed 1) a "spiral intensity profile" to detect colon cell location and size, and 2) an image segmentation algorithm that integrates multi-resolution, directional maximums-based edge detection with a path cost analysis technique. In the spiral intensity profile method, images (Fig. 7a) are enhanced (Fig. 7b), first by histogram equalization, and then by coherence enhancement diffusion (CED) [6], and then the colon cell centers are located by investigating the intensity profile along a spiral. As shown in Fig. 7c, by unwrapping pixels in a spiral to locate colon cell centers, colon cells of different sizes and cytoplasmic characteristics can be detected. Next, the results from the algorithm are used to initialize a segmentation algorithm, such as the level set segmentation algorithm [7],

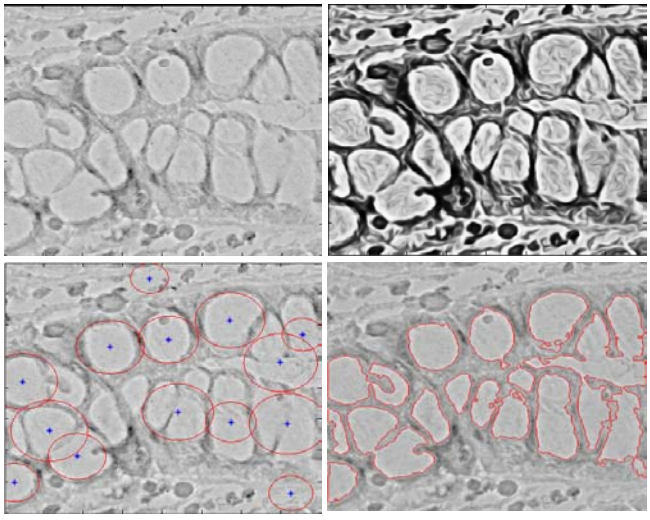

Fig. 7. a) Top left: original colon biopsy image; b) top right: image enhanced using histogram equalization and coherence enhancement diffusion; c) bottom left: colon cells then detected using a spiral intensity profile algorithm; d) bottom right: cell borders then extracted using a level set segmentation algorithm.

to accurately extract the border of each cell (Fig. 7d). In the "directional maximums and edge path cost analysis segmentation algorithm": 1) the image is prepared for segmentation using smoothing; 2) eight coding images are generated for directional maximums; 3) two profiling filters are used to obtain high and low edge density images; 4) the edge images are cleaned using skeletonization and spur removal techniques; 5) unconnected boundary ends are identified as nodes using spur detection; 6) path cost analysis is performed, and preferred paths are used to achieve continuous boundaries of the objects; and 7) objects with closed boundaries are segmented.

V. Summary

We have demonstrated that a valid panel of treatable biomarkers of risk for colorectal cancer can be detected in histologic slides of normal colon tissue (obtained in a simple, outpatient procedure somewhat comparable to a Pap smear) using traditional immunohistochemistry and a custom developed image analysis algorithm and program. We have also demonstrated promise that this approach can be improved and brought to clinical usefulness by upgrading our earlier methods with multiplexed, quantum dot immunohistochemistry and incorporating automated pattern recognition algorithms into our image analysis program. An example of the latter is the use of a spiral intensity profile to detect colon cell centers and approximate the sizes of the cells, and to initialize other segmentation algorithms, such as a level set segmentation algorithm. These advances hold promise for clinically valid, practical methods of assessing and managing risk for colorectal cancer.

References

[1] A. Jemal, R. Siegal, E. Ward, T. Murray, J. Xu , C. Smigal, and M. J. Thun, "Cancer Statistics, 2006," *Ca Cancer J. Clin.*, vol. 56, pp. 106-130, 2006.

[2] J. D. Potter , M. L. Slattery, R. M. Bostick, and S. M. Gapstur, "Colon cancer: a review of the epidemiology," *Epidemiol. Reviews,* vol. 15, pp. 499-545, 1993.

[3] G. E. Fraser, *Preventive Cardiology*. New York: Oxford University Press, 1986.

[4] J. D. Potter, "Colorectal cancer: molecules and populations," *J. Natl. Cancer Inst.,* vol. 91, pp. 916-32, 1999.

[5] K. W. Kinsler and B. Vogelstein, "Colorectal tumors," in *The Genetic Basis of Human Cancer*, 2nd ed., B. Vogelstein and K. W. Kinsler, Eds. New York: McGraw-Hill, 2002, pp. 583-612.

[6] J. Weickert, "A review of nonlinear diffusion filtering in scale-space theory in computer vision," in *Lecture Notes in Computer Science*, vol. 1252. Berlin: Springer-Verlag, 1997, pp. 3–28.

[7] Y. H. Tsai and S. Osher, "Level set methods in image science," in *2003 Proc. Intl. Conf. on Image Processing,* pp. 631-4.