



Published in final edited form as:

*Mod Pathol.* 2010 December ; 23(12): 1624–1633. doi:10.1038/modpathol.2010.161.

## Array Comparative Genomic Hybridization in Ulcerative Colitis Neoplasia: Single Non-Dysplastic Biopsies Distinguish Progressors from Non-Progressors

Mary P. Bronner<sup>1</sup>, Marek Skacel<sup>2</sup>, David A. Crispin<sup>3</sup>, Peter D. Hoff<sup>4</sup>, Mary J. Emond<sup>4</sup>, Lisa A. Lai<sup>3</sup>, Raymond R. Tubbs<sup>1</sup>, Peter S. Rabinovitch<sup>6</sup>, and Teresa A. Brentnall<sup>3</sup>

<sup>1</sup>Division of Pathology and Laboratory Medicine, Cleveland Clinic, Cleveland, Ohio, USA

<sup>2</sup>Dahl-Chase Pathology Associates, Bangor, Maine, USA

<sup>3</sup>Department of Medicine in the Division of Gastroenterology; University of Washington, Seattle, Washington, USA

<sup>4</sup>Department of Biostatistics, University of Washington, Seattle, Washington, USA

<sup>6</sup>Department of Pathology; University of Washington, Seattle, Washington, USA

### Abstract

Approximately 10% of ulcerative colitis patients develop colorectal neoplasia. At present, identification of this subset is markedly limited and necessitates lifelong colonoscopic surveillance for the entire ulcerative colitis population. Better risk markers are needed to focus surveillance onto the patients most likely to benefit. Using array-based comparative genomic hybridization, we analyzed single, *non-dysplastic* biopsies from three patient groups: ulcerative colitis progressors (n=9) with cancer or high-grade dysplasia at a mean distance of 18 cm from the analyzed site; ulcerative colitis nonprogressors (n=8) without dysplasia during long-term surveillance; and non-ulcerative colitis normal controls (n=2). Genomic DNA from fresh colonic epithelium purified from stroma was hybridized to 287 (low-density) and 4,342 (higher-density) feature bacterial artificial chromosome arrays. Sample-to-reference fluorescence ratios were calculated for individual chromosomal targets and globally across the genome. The low-density arrays yielded pronounced genomic gains and losses in 3 of 9 (33%) ulcerative colitis progressors but in none of the 10 control patients. Identical DNA samples analyzed on the higher density arrays, using a combination of global and individual high variance assessments, distinguished all 9 progressors from all 10 controls. These data confirm that genomic alterations in ulcerative colitis progressors are widespread, even involving single non-dysplastic biopsies far distant from neoplasia. They therefore show promise toward eliminating full colonoscopic surveillance with extensive biopsy sampling in the majority of ulcerative colitis patients.

Users may view, print, copy, download and text and data- mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Corresponding Author: Mary P. Bronner, M.D. Cleveland Clinic Department of Anatomic Pathology, L-25 9500 Euclid Ave. Cleveland, OH 44195 Ph: (216) 444-4833 Fax: (216) 445-6967 [bronnem@ccf.org](mailto:bronnem@ccf.org).

## Keywords

Ulcerative Colitis; Array Comparative Genomic Hybridization; Colonic Neoplasia

---

## Introduction

Long-term ulcerative colitis is characterized by chronic inflammatory injury to the colonic mucosa. A subset approximating 10% of these patients develop cumulative genomic damage leading to colorectal neoplastic progression (1-4). Currently detection of early, curable neoplasia requires life-long colonoscopic biopsy surveillance of the entire ulcerative colitis population, despite the knowledge that the great majority of these patients will never develop cancer. The large potential for sampling error is a major challenge. The surface area of the average colon ranges between 0.5 to 1.0 square meters relative to the tiny size of a biopsy (approximating 0.002% of the colonic mucosa) and the fact that most neoplasia is not visible by routine colonoscopy combine to create the problem. Accordingly a minimum of 33 jumbo, well-oriented random biopsies must be examined histologically to detect dysplasia with even 90% confidence (5), a number that is rarely achieved in clinical practice. Chromoendoscopy and other new endoscopic imaging techniques show promise for improving sampling error (6,7), but these technologies are not yet standard practice. Observer variability among pathologists and differentiation between inflammatory and neoplastic change are further considerable problems. Finally, at least half a million people in the United States have ulcerative colitis for whom lifelong colonoscopic examination involves considerable discomfort, high expense, and time and effort for patients and physicians alike. Unquestionably, better risk markers than dysplasia are needed.

Greater understanding of the molecular alterations in ulcerative colitis neoplastic progression may facilitate discovery of improved cancer risk markers. Previous reports have identified large fields of genomic damage in neoplastic progression in ulcerative colitis (1-3,5,8-15). Of great clinical importance, recent data even suggest that the abnormal molecular field may involve the entire colon. This might be anticipated from the diffuse chronic inflammatory injury thought to drive the process. Importantly, the field includes mucosal samples lacking histologic dysplasia that are far removed from morphologically neoplastic lesions (1-3,5,8-15). This offers great promise for reducing the serious problem of sampling error in surveillance. Such molecular markers could therefore focus individualized and more cost-effective surveillance onto the minority of ulcerative colitis patients who would most benefit. Of at least equal importance, such markers would allow the majority of ulcerative colitis patients who never develop cancer to safely forego difficult and unnecessary surveillance procedures.

While the potential of the genomic assays is exciting, they are technically difficult to perform. Furthermore, the distinction between ulcerative colitis progressors and nonprogressors using these markers remains imperfect. None of the proposed molecular markers have yet been validated prospectively or developed into high throughput tests amenable to clinical testing. To continue to explore and improve the discriminatory ability of genomic markers of cancer risk in ulcerative colitis, we now examine array-based

comparative genomic hybridization. This technology allows more comprehensive genome-wide scanning for chromosomal losses and gains over large numbers of targets for chromosomal gains and losses in comparison to earlier techniques (16,17). Discrimination of the genomic “signatures” in ulcerative colitis progressors relative to nonprogressors may provide a more sensitive and specific biomarker of cancer risk.

## Materials and Methods

### Patients

This research was approved by the Human Subject Division at the University of Washington and the Internal Review Board at the Cleveland Clinic. A total of 19 patients from the University of Washington were studied from three clinical groups: nine ulcerative colitis patients with cancer or high-grade dysplasia elsewhere in their colectomies (ulcerative colitis progressors), eight ulcerative colitis patients without dysplasia or flow cytometric DNA ploidy abnormalities on extensive biopsy sampling while in long-term colonoscopic biopsy surveillance for a minimum of 10 years (ulcerative colitis nonprogressors) and two non-ulcerative colitis control patients undergoing sigmoid colectomy for diverticulosis. Of the nine ulcerative colitis progressors, seven had adenocarcinoma and two had high-grade dysplasia as their most advanced neoplasia at colectomy. The colectomy specimens were extensively mapped histologically, examining >100 sections per colon. For the nonprogressors, an average of 44 jumbo, well-oriented biopsies was obtained at each ulcerative colitis surveillance colonoscopy. Colonoscopies were performed at 6 month to 3 year intervals depending on prior surveillance findings. The average age of the 17 ulcerative colitis patients (eight males, nine females) was 45 years, and that of the two non-ulcerative colitis controls (both males) was 54 years. The known duration of ulcerative colitis was a mean of 17.7 years (range: 4 to 35 years) for ulcerative colitis progressors, and 16.9 years (range: 10 to 25 years) for ulcerative colitis nonprogressors, respectively. Table 1 indicates the clinical parameters relative to the extent of neoplastic progression and the known clinical cancer risk markers in the nine ulcerative colitis progressor patients.

### Biopsy Samples

The samples of all progressors and two nonprogressors were obtained from densely mapped colectomy specimens. Those from the remaining six nonprogressors were from endoscopic biopsies from densely biopsied colonoscopies. All samples were rapidly frozen in cell culture minimal essential medium with 10% DMSO, as previously described (5). Tested colonic mucosal samples were split in half. One half was processed for comparative genomic hybridization per below. The other immediately adjacent half was routinely fixed, processed and sectioned to determine the histologic diagnoses, according to consensus criteria of the Inflammatory Disease/Dysplasia Morphology Study Group (DMSG) (18). All diagnoses were confirmed by one pathologist (MPB). The single histologically non-dysplastic biopsies were selected from within densely mapped regions of nondysplastic mucosa from the ulcerative colitis progressors. They were located at a mean distance of 18 cm (range: 2 to 79 cm; SD: 23.9 cm) from the most advanced neoplastic lesion (either carcinoma or high-grade dysplasia, respectively), and were located in the rectosigmoid colon (n=3), transverse colon (n=5) or right colon (n=1). The single non-dysplastic samples from

the nonprogressors were obtained from the left colon (n=5), right colon (n=1) and rectum (n=2). Both non-ulcerative colitis normal control samples were obtained from the sigmoid colon.

### Epithelial Isolation and DNA Extraction

Epithelial crypt isolation was performed using an ethylenediaminetetraacetic acid (EDTA) shake-off technique to isolate >90% pure colonic epithelial cells from mucosal stroma, as previously described (2). Whole genomic DNA was extracted using Genra Puregene kit according to manufacturer's instructions.

### Low Density Arrays

A commercially available genomic DNA microarray kit was utilized (GenoSensor Array 300; Vysis, Downers Grove, IL), containing 287 oncologically representative genomic DNA targets from P1, PAC, and bacterial artificial chromosome libraries. Aliquots of sample and normal pooled reference DNA were DNase digested to generate 50- to 800-base pair (bp) fragments, labeled by random priming reaction (Random Priming Reaction kit; Vysis) with Cy 3-dCTP or Cy 5-dCTP (Perkin Elmer, Boston, MA), respectively, and hybridized in Microarray Hybridization Buffer (Vysis) containing Cot-1 DNA at 80 °C for 10 minutes. Hybridization to the array slide at 37 °C for 60 hours in a humidified coplin jar with 2.0 mL 50% formamide and 2X standard saline citrate (SSC) was followed by washes and counterstaining with 20 µL of 4'6-diamidinophenylindole dihydrochloride (DAPI) IV solution (Vysis). The data obtained from the low-density arrays were analyzed using the built-in software of the GenoSensor™ Reader system (Vysis). Based on average mean fluorescence intensity across the array, this software calculates the correlation coefficients among the triplicate spots for each target and displays the fluorescence ratios between the sample and reference DNA when they fulfill all quality control criteria set by the manufacturer.

### Higher Density Arrays

DNA was digested with Dpn II and applied to a MinElute Reaction Clean-Up column (QIAGEN). Random amplification and labeling steps were performed according to the method of Lieb with modifications (19). Briefly, a primer (Primer A; 5'GTTTCCCAGTCACGATC NNNNNNNNN3') containing a degenerate 3' end and a specific 5' end was randomly annealed to DNA templates and extended with a T7 polymerase (Sequenase Version 2.0, USB, Cleveland, OH). Then, a Cy3- or Cy5 labeled primer (5'GTTTCCCAGTCACGATC; Operon, Alameda, CA) specific to the 5' end of Primer A was used to PCR amplify the templates generated in the first reaction. PCR reactions were then applied to a QIAquick PCR Purification column (QIAGEN), blocking agents added (50 µg human Cot-1 DNA and 100 µg yeast tRNA; Invitrogen, Carlsbad, CA) and concentrated with a Microcon 10 filter (Millipore, Bedford, MA).

The bacterial artificial chromosome array was manufactured at the Fred Hutchinson Cancer Research Center array facility and comprised 4,342 clones (20), spotted in triplicate. The clones are a subset of those described in the bacterial artificial chromosome Resource Consortium (21), augmented with clones containing genes relevant to tumorigenesis and

clones isolated from single-copy regions near the end of most chromosome arms (22). The genomic inserts of the bacterial artificial chromosomes average 166 kb. Some sequence has been determined from each clone (e.g. end sequence read(s), PCR-confirmed sequence-tagged site (STS) or gene, or the clone's entire sequence), allowing each clone to be positioned (at the midpoint of each clone) precisely within the human draft sequence assembled as of April 2003 (23), <http://genome.ucsc.edu/>). Clones that hybridized by FISH to multiple chromosomal locations or to a cytogenetic position inconsistent with a position in the draft sequence assembly, were excluded from analysis, yielding 4,153 clones used in the final analysis (21). These clones are distributed across the genome at a median spacing of 411 kb when the repetitive regions around centromeres and on the short arms of acrocentric chromosomes are excluded (25<sup>th</sup> percentile, 161 kb; 75<sup>th</sup> percentile, 869 kb; mean, 675 kb; 3 intervals exceed 5 MB). The locations of 94% of the 4,153 clones have been verified by FISH. The clones used in our analyses and details describing the coordinates are available from <http://www.fhcr.org/labs/trask/arraycomparative> genomic hybridization. The clones are available from CHORI [<http://bacpac.chori.org/>].

To construct the array, bacterial artificial chromosome DNA was extracted from 1.5-ml cultures using a modified Qiagen R.E.A.L. Prep96 plasmid kit (QIAGEN). Purified bacterial artificial chromosome DNA was then amplified using ligation-mediated PCR (23) and cleaned using the Multiscreen-PCR filtration system (Millipore) on a Beckman Coulter Biomek FX automated liquid handling system (BeckmanCoulter, Fullerton, CA). Purified products were transferred to 384-well plates, air-dried, and then resuspended in 3xSSC. Each amplified/purified bacterial artificial chromosome was then mechanically spotted in triplicate onto poly-l-lysine-coated slides using a GeneMachines Omnigrid Arrayer (GeneMachines, San Carlos, CA).

Tumor and reference pooled normal DNA labeled with Cy5 and Cy3, respectively, were combined and briefly dried in a speedvac. The samples were then resuspended in hybridization mix (50% formamide, 10% dextran sulfate, 3xSSC, 1.5% SDS). Immediately following resuspension, samples were heated to 70°C for 10 min followed by 1-hr incubation at 37°C to allow blocking of repetitive sequences by human Cot-1 DNA. After incubation, the sample was applied to the microarray under a coverslip and hybridized at 37°C for 18 hrs. Following hybridization, the arrays were then washed in 50% formamide, 2xSSC, pH 7.0, for 20 minutes at 45°C followed by a 7-min wash at room temperature in 0.1M NaPO<sub>4</sub>, pH 8.0, 0.1% NP40. After the washes, the arrays were dried by centrifugation at 600 rpm for 5 min.

The higher density arrays were scanned using the GenePix 4000B scanner (Axon Instruments Inc, Union City, CA.), and the fluorescence data were processed using GenePix 3.0 image analysis software (Axon Instruments, Inc.). For each spot,  $\log_2(\text{ratio}) = \log_2(\text{Cy5}/\text{Cy3})$  was calculated, where Cy5 and Cy3 refer to the median foreground fluorescent signals of the tumor and reference DNA (respectively). The  $\log_2$  ratios on each array were normalized and corrected for intensity-based location adjustment using a block-level loess algorithm (24).

## Reference DNA

Two control DNA pools were used for array comparative genomic hybridization analysis. The male control and female control pools each contained DNA from 15 cytogenetically normal individuals. For procedural quality control, all analyses were performed as sex-mismatch hybridizations. This allowed determinations of chromosome X and Y copy number as an internal reference standard.

## Statistical Analysis

The low-density array data was analyzed in respect to overall extent of genomic abnormalities per sample as identified by the built-in GenoSensor software. The frequencies of genomic gains and losses in progressors, nonprogressors and non-ulcerative colitis controls were compared using descriptive statistics. The higher-density array data were analyzed in respect to an overall extent of genomic aberrations as well as to specific abnormalities differing between progressors and nonprogressors. Copy number alterations were independently verified using comparative genomic hybridization-Miner (25) with a 3 SNP moving window and false discovery rate set at 0.01. Copy number estimates were manually curated for events with an average  $\log_2$  intensity ratio of  $> 0.26$  (gains) or  $< -0.32$  (losses). As a measure of the overall amount of copy number aberration over the entire genome, the following procedure was performed: An absolute value of the  $\log_2$  copy number ratio for each locus was calculated (normal ratio is 1, normal  $\log_2$  ratio is 0). The mean of the most extreme 5% of these absolute values was defined to be the “extreme deviation score” for each patient. The 5% value was chosen arbitrarily, analogous to common assumption that the most extreme 5% of a measurement on a population is most likely to be representative of biological abnormalities. In order to identify specific aberrations discriminating ulcerative colitis progressors from nonprogressors, we restricted our search to the subset of bacterial artificial chromosomes showing a significantly higher variance in response than the overall sample variance in response. This selection was independent of the progressor/nonprogressor status of the cases. One-sided chi-square test of equal variance and Bonferroni correction for multiple comparisons were used in this analysis. The Wilcoxon rank-sum test was then used to compare responses of progressors and nonprogressors for each of the identified candidate bacterial artificial chromosome clones.

## Results

### Low Density Arrays

Using the low-density arrays, marked genomic alteration detected at 16 or more of the 287 loci on the array was identified in three of nine (33%) biopsies from ulcerative colitis progressors (mean: 18.3 altered loci; range: 17-19; SD: 1.2). Only rare genomic changes detected at  $<5$  loci were observed in the remaining six ulcerative colitis progressors (mean: 2.5 loci; range: 0-4; SD: 1.4), as well as all eight ulcerative colitis nonprogressors (mean: 1.3 loci; range: 0-3; SD: 1.1) and both non-ulcerative colitis controls (mean: 1.8; range 0-5; SD: 1.9). The absence of alterations in any of the ten controls, suggests good specificity. However, since two-thirds of the ulcerative colitis progressors went undetected, the low-

density array provided only limited sensitivity. Larger arrays were then studied to assess whether this might improve sensitivity.

### Higher Density Arrays

Visual inspection of the 4,153 effective element bacterial artificial chromosome arrays showed patterns suggestive of increased chromosomal gains and losses (Figure 1C) and interstitial deletions (Figure 1D) in ulcerative colitis progressors, as compared to normal colon controls (Figure 1A) or ulcerative colitis nonprogressors (Figure 1B). To create a quantitative measure of global copy number aberration throughout the entire genome, an absolute value of the  $\log_2$  copy number ratio for each bacterial artificial chromosome target on the array was calculated. Using the mean of the most extreme 5% of these absolute values, an “extreme deviation score” was calculated for each patient. In addition, as another independent metric to the deviation score, standard deviations of the results across all measurements were also calculated for each patient. Both analyses demonstrated increasing genomic “noisiness” for the ulcerative colitis progressors compared to nonprogressors as demonstrated in Figure 2 A and B, respectively. With the exception of two ulcerative colitis progressors (P1 and P2), who showed similar degrees of alteration as nonprogressors, the remaining seven progressors (78%) displayed consistently greater extreme deviation scores and standard deviations than any of the nonprogressors. Notably, three especially unstable progressors were identified using both approaches and are shown in the right upper end of the both curves (P7, P8 and P9). These three patients were also the only three detected as markedly unstable using the low-density arrays.

Furthermore, after obtaining a global “signature” of genomic aberrations for each patient, we assessed whether the presence of aberrations at particular chromosomal locations could discriminate among the two patient groups. We were interested in identifying chromosomal targets or combinations of targets for which all or most progressors have in common that are absent in the majority of nonprogressors. Individual target variance analysis confirmed the presence of a distinct pattern of genomic alteration in the ulcerative colitis progressors and showed an improved separation of these samples from nonprogressors. While the majority of aberrations detected genome-wide appeared random and did not discriminate between the patient groups, a smaller subset of changes showing the greatest variance were significantly more common in progressors. Using a one-sided chi-square test of equal variance with a Bonferroni-corrected significance level of  $\alpha=0.05/4,153$  a total of 40 candidate bacterial artificial chromosomal targets were identified as having unusually high variance and were selected as candidate discriminators of progressors versus nonprogressors. Wilcoxon rank-sum tests and t-tests comparing responses of progressors to those of nonprogressors were performed for each candidate target separately. Eleven of the targets exhibited both t-statistics and rank-sum statistics that were significant at the nominal 0.05 level of significance, and are given in Table 2. All eleven of these targets demonstrated genomic loss in the ulcerative colitis progressors.

To provide results that are corrected for multiple testing, a Bonferroni correction was applied to maintain an overall false positive error rate of 0.05 in declaring significant differences between progressors and nonprogressors at particular chromosomal sites. With

this method, a significant difference is declared only if the p-value is less than  $0.05/40 = 0.00125$ . This is the case for bacterial artificial chromosome targets 3905, 3906 and 3907 (RP11-7H17, RP11-196B3 and RP11-315M18), which all map to the distal end of the long arm of chromosome 18, (18q23), as shown in Table 2. This clustering, which was independent of selection process, provides further confirmation of the statistical significance of these results. The differences for bacterial artificial chromosome 944 (RP11-460I19, maps to 4p terminus) are of borderline significance by this criterion.

It should be noted that failure to show significance after correction for multiple testing for most of the eleven sites in Table 2 does not rule out the possibility that these sites have biological significance; it only means we were unable to show significance after correction for multiple testing. This can result either from lack of biological significance or from lack of statistical power to identify differences with few patients and large numbers of possible predictors. The (normalized) log-ratios among all of the patients for the three 18q sites 3905, 3906 and 3907 are shown in Figure 3. As can be seen,  $\log_2$ -ratios from these three sites separate the progressors from the nonprogressors except for one case (P3). Except for this one case, the remaining progressors all have negative values that are large when compared to the mean and standard deviation of responses across all sites (mean of  $-0.001$ , standard deviation of  $0.11$ ), suggesting chromosomal loss at the locations corresponding to these sites. In contrast, the distributions of the  $\log_2$ -ratios among the nonprogressors are centered on zero and are relatively small in magnitude, suggesting there is no relative chromosomal gain or loss. This suggests that progressors can generally be characterized by chromosomal loss on the distal tip of chromosome 18q.

## Discussion

It is hypothesized that genomic alterations leading to ulcerative colitis neoplasia result from long-standing chronic inflammatory mucosal injury. However, only an estimated 10% minority of the ulcerative colitis population ever develops neoplasia. As such, improved markers to identify this at risk group are greatly needed. This study identifies promising molecular markers using array-based comparative genomic hybridization.

Discoveries in the last several years have advanced our understanding of the molecular biology of ulcerative colitis. Widespread genomic alterations have been demonstrated in ulcerative colitis progressors with neoplasia, but not nonprogressors, using assays of telomere shortening and anaphase bridge development (2), DNA-fingerprinting by arbitrarily primed PCR (13-15), preclonal chromosomal gains and losses by fluorescence in situ hybridization (FISH) (1,2,28), metaphase-comparative genomic hybridization (26-28), and estrogen receptor gene promoter hypermethylation (29).

To our knowledge, this is the first report of array comparative genomic hybridization in ulcerative colitis. Our goal was to interrogate the genome broadly to: 1) identify specific alterations or patterns of chromosomal alterations that distinguish ulcerative colitis progressors from nonprogressors, and 2) to better understand the early events that lead to neoplastic progression in ulcerative colitis by analyzing mucosa without dysplasia. Array technology permits high resolution genomic scanning for chromosomal losses and gains



over many thousands of genomic targets. The much higher resolution afforded by array technology is a major advance over earlier comparative genomic hybridization technology using metaphase chromosomes (26-28). The higher density 4,298 feature BAC array used in this study has a mean intermarker spacing of 675 kilobase in comparison to the approximate five megabase resolution of metaphase chromosomes. Metaphase-comparative genomic hybridization also necessitates considerable cytogenetic expertise for the morphologic evaluation of hybridization signals along physical chromosomes rather than the automated signal detection of known chromosomal sites spotted onto systematic arrays. Regardless of the differences in these technologies, it is encouraging that a few differential chromosomal sites identified by earlier metaphase-comparative genomic hybridization were also observed in our array results, specifically chromosomal losses on 15, 5q, and 18q (26-28).

An initial low-resolution array (287 clones) provided proof of principle that a greater degree of clonal genomic alteration was in fact present in non-dysplastic mucosa from 33% the progressors (three of nine) in contrast to none of the eight nonprogressors or two normal controls. The low-density array data additionally provided preliminary data regarding locations of highest frequency genomic alterations. As anticipated, subsequent analysis of the identical DNA samples, using markedly higher-density arrays (4,153 clones), demonstrated this phenomenon on a much larger scale, distinguishing all nine progressors from all of the eight nonprogressors and two normal controls, when analyzed by a combination of statistical measures.

In addition, analysis of the higher-density arrays identified specific and unique genomic losses associated with neoplastic progression in dysplasia-free colonic samples from ulcerative colitis progressors. Of importance, the biopsy samples were obtained far distant from their respective tumors, highlighting the widespread nature of the genomic alterations within the colon. However, given the 4,153 clones analyzed, it is not unlikely that due to chance alone a few targets in the progressors will all have gains or losses below some threshold and nonprogressors above, or vice-versa, which could thus falsely discriminate perfectly among the groups. If the responses are not large in magnitude, then it is likely that the group differences are a result of experimental noise and not due to consistent chromosomal gains or losses. To eliminate these from consideration, we restricted our search for potential discriminating targets to the subset of targets having a significantly higher variance in response than the overall sample variance in response. This selection was independent of the progressor/nonprogressor status of the cases. The reasoning behind this selection criterion is as follows: for most genomic locations, most of the 19 cases do not have genomic abnormalities, and have array response intensity deviations that are relatively small in magnitude and centered around zero. In contrast, if many cases have abnormalities at a location corresponding to a given clone, then many of the 19 responses will be large in magnitude and the sample variance for that target will be large. By selecting for consideration only those bacterial artificial chromosomes with high variance, we are selecting targets for which there is evidence of genomic aberration across some of the 19 cases, although no consideration of group membership is taken into account, eliminating bias.

While a portion of the chromosomal gains and losses were seen in both progressors and nonprogressors, most of them were random and not consistent between biopsies. However, a smaller subset of changes with great variance was non-random, as shown by t-tests comparing their distribution among progressors and nonprogressors (Table 2 lists the eleven most significant). Such non-random losses of several genomic regions were observed in progressors, with the most powerful discriminator being a chromosomal loss on RP11-196B3, which maps to the distal tip of the 18q arm (18q23). Two adjacent and overlapping targets at 18q were also highly discriminant. The 18q23 loss was detected in eight of nine (89%) progressors and was not seen in any of the eight nonprogressors or two normal controls. Other than several sequence tagged sites, there is only one known gene in this area, named “nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1” (NFATC1), according to the NCBI clone registry. Of interest, this transcription factor initially identified as an important modulator of T-cell cytokine production (30,31) has been found to have increasing roles outside the immune system. A recent study by Duque and colleagues documents its importance in COX-2 regulation in colorectal carcinoma cell lines (31). COX-2 upregulation is well-described in colorectal tumorigenesis, both sporadic and ulcerative colitis forms (12).

It is interesting that this 18q as well as other highly discriminant loci (5p15.33 and 10q26.3) by array comparative genomic hybridization are located in subtelomeric chromosomal regions. This finding further implicates a chromosomal bridge-fusion-breakage mechanism in relation to telomere shortening in ulcerative colitis neoplastic progression (1,2) and in other cancers (33,34). Subtelomeric regions may be inherently more prone to breakage after telomere attrition-mediated end-to-end fusions with breakage at anaphase.

It is also noteworthy that the highest variance non-random chromosomal alterations were all genomic losses (Table 2). This implicates loss of tumor suppressor genes as an important mechanism of the molecular pathogenesis of ulcerative colitis neoplasia.

These array comparative genomic hybridization data provide evidence that genomic changes associated with neoplastic progression in ulcerative colitis are sufficiently advanced to be clonal, which is intriguing since they are present in non-dysplastic single random biopsies from ulcerative colitis progressors. That they are clonal derives from the knowledge that the detection threshold for comparative genomic hybridization gains or losses requires the majority of cells to demonstrate a given alteration (34).

Three progressors failed to simultaneously demonstrate all three statistical measures of array comparative genomic hybridization alteration. Specifically, one of the tested ulcerative colitis progressors (P3) in our small series of nine patients did not exhibit the 18q alterations (Figure 3), and an additional two ulcerative colitis progressors (P1 and P2) failed to demonstrate the random global alterations assessed by extreme 5th percentile deviation and standard deviation analyses (Figures 2A and 2B). This could be related to the still limited resolution of ~ 1 MB afforded by even the higher density 4,298 bacterial artificial chromosome arrays utilized in this study. The improved sensitivity gained by the higher density array supports this hypothesis. Improved resolution afforded by SNP, oligo and tiling or submegabase-resolution tiling (SMRT) arrays may improve sensitivity even further.

These three patients' results may also relate to sampling error within the colon, namely that more than one biopsy may be necessary to identify progressors and nonprogressors. Future mapping studies of the distribution of CGH alterations throughout the colon of both progressors and nonprogressors will be important.

A final potential explanation for why three ulcerative colitis progressors were not detected by all three statistical measures in this study could be that these patients differed clinically. This was true for the two that lacked global genomic instability, namely P1 and P2 (Table 1). Progressor P2 had incident high-grade dysplasia that was detected during a prolonged 20 years of surveillance endoscopy, in comparison to the much more advanced prevalent adenocarcinomas in the majority of the remaining progressors. Further, progressor P1 was the only ulcerative colitis progressor without ulcerative pancolitis. Instead, progressor P1 had disease limited to the rectum rather than the pancolonic ulcerative colitis in the remaining progressors. Numerous studies document that ulcerative proctitis appears to be a different disease with limited to no increased cancer risk. Surveillance is, in fact, not recommended for the ulcerative proctitis variant of ulcerative colitis (35-37). The biopsy analyzed by comparative genomic hybridization for patient P1 derived from the unaffected sigmoid colon without changes of ulcerative colitis. It is possible that patient P1, who was also 74 years of age, actually developed a sporadic-type rectal cancer rather than an ulcerative colitis cancer. This remains speculation but this patient was certainly clinically distinct from the remaining ulcerative colitis progressors. Thus, it is possible that these clinical differences may explain at least part of the variant array comparative genomic hybridization results for these three ulcerative colitis progressors that were not detected by all three statistical measures.

The fact that far removed single biopsies lacking dysplasia in our small series of 17 ulcerative colitis cases were able to distinguish progressors and nonprogressors is remarkable. From a clinical perspective, this adds further credence to the concept that full colonoscopy with extensive biopsy sampling, currently the standard of care, could possibly be eliminated in the future. This follows from the finding that even single biopsies may harbor the genomic signature identifying ulcerative colitis progressors that could focus surveillance efforts onto the subset most likely to benefit. Natural history data on the development of genomic alterations during ulcerative colitis progression remain to be determined, as the current study only evaluated biopsies only at the time of synchronous high-grade dysplasia or carcinoma. Longitudinal and prospective validation studies are underway.

In summary, array comparative genomic hybridization is a powerful discovery technique for genomic markers of ulcerative colitis cancer risk to better identify the minority subset of patients in need of intensive colonoscopic surveillance. By sequentially utilizing low- and high-resolution arrays, we demonstrate that the power of this approach increases with the increasing array target density. Comparing the patterns of genomic alteration in ulcerative colitis progressors with nonprogressors, random and specific alterations identify progressors with 100% sensitivity and specificity in our small series of 17 ulcerative colitis patients. These results were achieved using only single non-dysplastic biopsies far distant from the patients' neoplasms. Of the changes identified, loss at the subtelomeric region of

chromosome 18q appears to be the most powerful marker, along with several other subtelomeric losses, further implicating the telomere shortening mechanism with bridge-breakage-fusion cycles previously identified in ulcerative colitis neoplastic progression. These array data represent an important additional step toward understanding of molecular events in neoplastic development in ulcerative colitis, and further demonstrate the great promise of molecular markers on even single biopsy samples to advance ulcerative colitis cancer surveillance.

## Acknowledgments

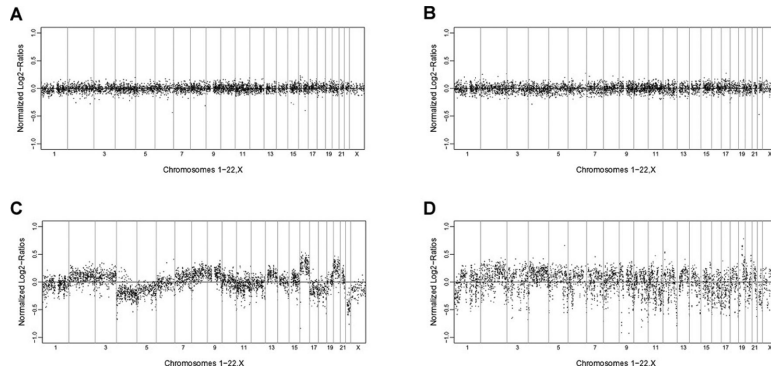
Grant Support: NIH R01 CA068124

## References

- O'Sullivan J, Bronner MP, Brentnall TA, Finley JC, Shen W-T, Emerson S, Emond MJ, Gollahon KA, Moskovitz AH, Crispin DA, Potter JD, Rabinovitch PS. Chromosomal alteration is related to telomere shortening in a human preneoplastic disease (ulcerative colitis). *Nat Genet.* 2002; 32:280–284. [PubMed: 12355086]
- Bronner MP, O'Sullivan JN, Rabinovitch PS, Crispin DA, Chen L, Emond MJ, Rubin CE, Brentnall TA. Genomic biomarkers on distant non-dysplastic biopsies distinguish UC progressors with neoplasia from dysplasia-free nonprogressors. *Am J Pathol.* 2008; 173:1853–1860. [PubMed: 18988798]
- Brentnall TA, Rubin CE, Crispin DA, Stevens A, Batchelor RH, Haggitt RC, Bronner MP, Evans JP, Boland CR, Bilir N, Rabinovitch PS. A germline substitution in the MSH2 gene is associated with neoplasia in UC. *Gastroenterology.* 1995; 109:151–155. [PubMed: 7797014]
- Brentnall TA. Molecular underpinnings of cancer in UC. *Curr Opin Gastroenterol.* 2003; 19:64–68. [PubMed: 15699896]
- Rubin CE, Haggitt RC, Burner GC, Brentnall TA, Stevens A, Levine DS, Dean PJ, Kimmey MB, Perera DR, Rabinovitch PS. DNA aneuploidy in colonic biopsies predicts future development of dysplasia in UC. *Gastroenterology.* 1992; 103:1611–1620. [PubMed: 1426881]
- Rutter MD, Saunders BP, Schofield G, Forbers A, Price AB, Talbot IC. Pancolonial indigo carmine dye spraying for the detection of dysplasia in UC. *Gut.* 2004; 53:256–260. [PubMed: 14724160]
- Kiesslich R, Neurath MF. Chromoendoscopy: an evolving standard in surveillance for UC. *Inflamm Bowel Dis.* 2004; 10:695–696. [PubMed: 15472537]
- Burner GC, Rabinovitch PS, Haggitt RC, Crispin DA, Brentnall TA, Kolli VR, Stevens A, Rubin CE. Neoplastic progression in UC: Histology, DNA content, and loss of a p53 allele. *Gastroenterology.* 1992; 103:1602–1610. [PubMed: 1358743]
- Brentnall TA, Crispin DA, Rabinovitch PS, Haggitt RC, Rubin CE, Stevens AC, Burner GC. Mutations in the p53 gene: an early marker of neoplastic progression in UC. *Gastroenterology.* 1994; 107:369–378. [PubMed: 8039614]
- Bronner MP, Culin C, Reed JC, Furth EE. The bcl-2 proto-oncogene and the gastrointestinal epithelial tumor progression model. *Am J Pathol.* 1995; 146:20–26. [PubMed: 7856728]
- Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. Accelerated age-related CpG island methylation in UC. *Cancer Res.* 2001; 61:3573–3577. [PubMed: 11325821]
- Agoff SN, Brentnall TA, Crispin DA, Taylor SL, Raaka S, Haggitt RC, Reed MW, Afonina IA, Rabinovitch PR, Stevens AC, Feng Z, Bronner MP. The role of COX-2 in UC-associated neoplasia. *Am J Pathol.* 2000; 157:737–745. [PubMed: 10980113]
- Chen R, Rabinovitch PS, Crispin DA, Emond MJ, Bronner MP, Brentnall TA. The initiation of colon cancer in a chronic inflammatory setting. *Carcinogenesis.* 2005; 26:1513–1519. [PubMed: 15860506]

14. Chen R, Bronner MP, Crispin DA, Rabinovitch PS, Brentnall TA. Characterization of genomic instability in UC neoplasia leads to discovery of putative tumor suppressor regions. *Cancer Genet Cytogenet.* 2005; 162:99–106. [PubMed: 16213356]
15. Chen R, Rabinovitch PS, Crispin DA, Emond MJ, Koprowicz KM, Bronner MP, Brentnall TA. DNA fingerprinting abnormalities can distinguish UC patients with dysplasia and cancer from those who are dysplasia/cancer-free. *Am J Pathol.* 2003; 162:665–672. [PubMed: 12547724]
16. Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer.* 1997; 20:399–407. [PubMed: 9408757]
17. Pinkel D, Seagraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nature Genet.* 1998; 20:207–211. [PubMed: 9771718]
18. Riddell RH, Goldman H, Ransohoff DF, Appelman HD, Fenoglio CM, Haggitt RC, Ahren C, Correa P, Hamilton SR, Morson BC. Dysplasia in inflammatory bowel disease: standardized classification with provisional clinical applications. *Human Pathol.* 1983; 14:931–968. [PubMed: 6629368]
19. Lieb J, Liu X, Botstein D, Brown P. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat Genet.* 2001; 28:327–334. [erratum appears in *Nat Genet* 29:100; 2001]. [PubMed: 11455386]
20. Loo LW, Grove DI, Williams EM, Neal CL, Cousens LA, Schubert EL, Holcomb IN, Massa HF, Glogovac J, Li CI, Malone KE, Daling JR, Delrow JJ, Trask BJ, Hsu L, Porter PL. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. *Cancer Res.* 2004; 64:8541–8549. [PubMed: 15574760]
21. Cheung V, Nowak N, Jang W, Kirsch I, Zhao S, Chen X, Furey T, Kim U, Kuo W, Olivier M, Conroy J, Kasprzyk A, Massa H, Yonescu R, Sait S, Thoreen C, Snijders A, Lemyre E, Bailey J, Bruzel A, Burrill W, Clegg S, Collins S, Dhami P, Friedman C, Han C, Herrick S, Lee J, Ligon A, Lowry S, Morley M, Narasimhan S, Osoegawa K, Peng Z, Plajzer-Frick I, Quade BJ, Scott D, Sirotkin K, Thorpe A, Gray J, Hudson J, Pinkel D, Ried T, Rowen L, Shen-Ong G, Strausberg R, Birney E, Callen D, Cheng J, Cox D, Doggett N, Carter N, Eichler E, Haussler D, Korenberg J, Morton C, Albertson D, Schuler G, de Jong P, Trask B. Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature.* 2001; 409:953–958. [PubMed: 11237021]
22. Knight SJ, Lese CM, Precht KS, Kuc J, Ning Y, Lucas S, Regan R, Brenan M, Nicod A, Lawrie NM, Cardy DL, Nguyen H, Hudson TJ, Riethman HC, Ledbetter DH, Flint J. An optimized set of human telomere clones for studying telomere integrity and architecture. *Am J Hum Genet.* 2000; 67:320–332. [PubMed: 10869233]
23. Kent W, Sugnet C, Furey T, Roskin K, Pringle T, Zahler A, Haussler D. The human genome browser at UCSC. *Genome Res.* 2002; 12:996–1006. [PubMed: 12045153]
24. Yang YH, Dudoit S, Luu P, Speed TP. Normalization for cDNA microarray: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* 2002; 30:e15. [PubMed: 11842121]
25. Wang P, Kim Y, Pollack J, Narasimhan B, Tibshirani R. A method for calling gains and losses in array CGH data. *Biostatistics.* 2005; 6:45–58. [PubMed: 15618527]
26. Kupka S, Schröder K, Porschen R, Borchard F, Gregor M, Blin N, Holzmann K. Comparative genomic hybridization analysis of chromosomal alterations in patients with long-standing UC. *Int J Oncol.* 2001; 19:489–494. [PubMed: 11494025]
27. Willenbacher RF, Aust DE, Chang CG, Zelman SJ, Ferrell LD, Moore DH 2nd, Waldman FM. Genomic instability is an early event during the progression pathway of ulcerative-colitis-related neoplasia. *Am J Pathol.* 1999; 154:1825–1830. [PubMed: 10362807]
28. Willenbacher RF, Zelman SJ, Ferrell LD, Moore DH 2nd, Waldman FM. Chromosomal alterations in ulcerative colitis-related neoplastic progression. *Gastroenterology.* 1997; 113:791–801. [PubMed: 9287970]
29. Fujii S, Tominaga K, Kitajima K, Takeda J, Kusaka T, Fujita M, Ichikawa K, Tomita S, Ohkura Y, Ono Y, Imura J, Chiba T, Fujimori T. Methylation of the oestrogen receptor gene in non-

- neoplastic epithelium as a marker of colorectal neoplasia risk in longstanding and extensive UC. *Gut*. 2005; 54:1287–1292. [PubMed: 15870230]
30. Serfling E, Berberich-Siebelt F, Chuvpilo S, Jankevics E, Klein-Hessling S, Twardzik T, Avots A. The role of NF-AT transcription factors in T cell activation and differentiation. *Biochim Biophys Acta*. 2000; 1498:1–18. [PubMed: 11042346]
  31. Rao A, Luo C, Hogan PG. Transcription factors of the NFAT family: regulation and function. *Annu Rev Immunol*. 1997; 15:707–747. [PubMed: 9143705]
  32. Duque J, Fresno M, Iniguez MA. Expression and function of the nuclear factor of activated T cells in colon carcinoma cells: involvement in the regulation of cyclooxygenase-2. *J Biol Chem*. 2005; 280:8686–8693. [PubMed: 15632146]
  33. Soler D, Genesca A, Arnedo G, Egozcue J, Tusell L. Telomere dysfunction drives chromosomal instability in human mammary epithelial cells. *Genes Chromosomes Cancer*. 2005; 44:339–350. [PubMed: 16052508]
  34. Lim G, Karaskova J, Beheshti B, Vukovic B, Bayani J, Selvarajah S, Watson SK, Lam WL, Zielenska M, Squire JA. An integrated Mband and submegabase resolution tiling set (SMRT) CGH array analysis of focal amplification, microdeletions, and ladder structures consistent with breakage–fusion–bridge cycle events in osteosarcoma. *Genes Chromosomes Cancer*. 2005; 42:392–403. [PubMed: 15660435]
  35. Kallioniemi OP, Kallioniemi A, Piper J, Isola J, Waldman FM, Gray JW, Pinkel D. Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors. *Genes Chromosomes Cancer*. 1994; 10:231–243. [PubMed: 7522536]
  36. Siproudhis L, Vilotte J, Bonfils S, Mignon M. Idiopathic ulcerative proctitis. Clinical presentation and endoscopic outcome. *Gastroenterol Clin Biol*. 1991; 15:315–321. [PubMed: 2060743]
  37. Ghirardi M, Nascimbeni R, Mariani PP, Di Fabio F, Salerni B. Course and natural history of idiopathic ulcerative proctitis in adults. *Ann Ital Chir*. 2002; 73:155–159. [PubMed: 12197289]
  38. Ekobom A, Helmick C, Zack M, Adami HO. Ulcerative colitis and colorectal cancer. A population-based study. *N Engl J Med*. 1990; 323:1228–1233. [PubMed: 2215606]



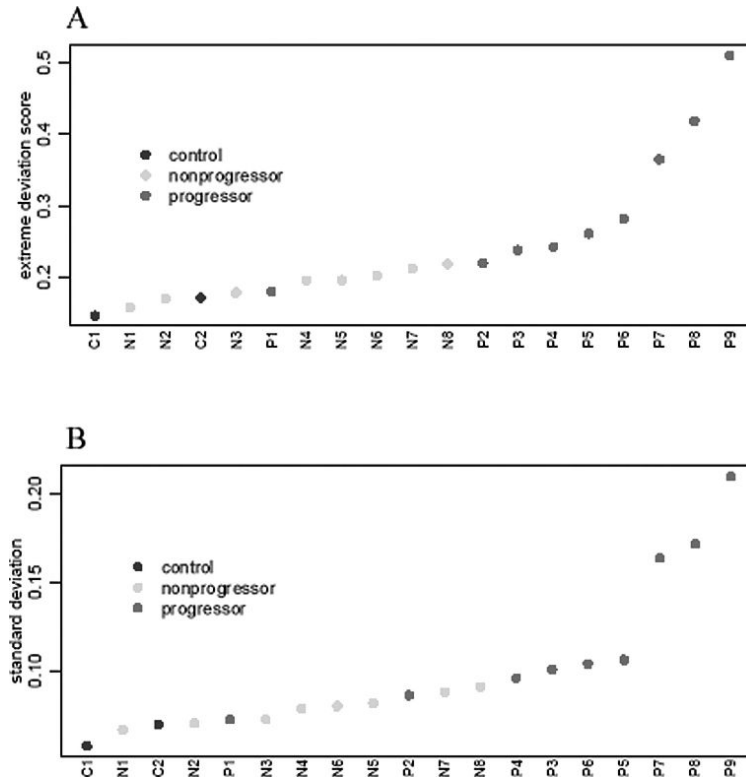
**Figure 1.** Higher density bacterial artificial chromosome array whole genome log<sub>2</sub>-ratios plotted across all chromosomes consecutively from 1 through 22 plus the X chromosome. A) normal control non-ulcerative colitis patient “C1”; B) ulcerative colitis nonprogressor patient “N1”; C) ulcerative colitis progressor patient “P8”; and D) ulcerative colitis progressor patient “P9”. These four representative patient results illustrate prominent chromosomal gains and losses (C) or interstitial deletions (D) in ulcerative colitis progressors in contrast to the minimal changes in nonprogressor (B) and normal controls (A).

Author Manuscript

Author Manuscript

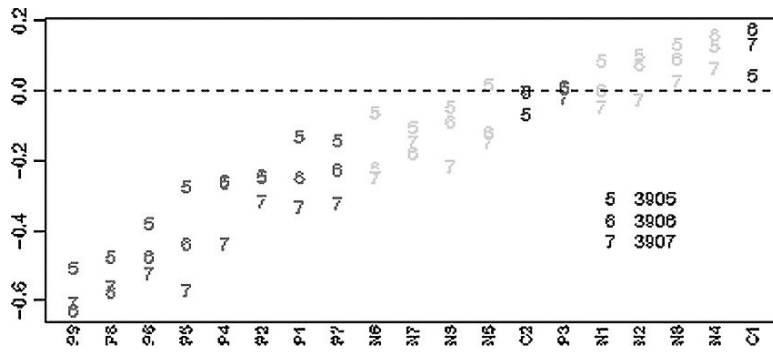
Author Manuscript

Author Manuscript



**Figure 2.** Extreme deviation score (2A) and standard deviation (2B) plots representing the absolute value of the  $\log_2$  copy number ratio for the case-to-reference ratio at each locus (normal ratio is 1, normal  $\log_2$  ratio is 0). The mean of the most extreme 5% of these absolute values was defined as the “extreme deviation score” for each patient. The standard deviation examines standard deviation results across all measurements. Other than for progressors P1 and P2, the remaining 78% of progressors displayed greater genomic deviation relative to ulcerative colitis nonprogressor and normal non-ulcerative colitis control patients. P=ulcerative colitis progressor; N=ulcerative colitis nonprogressor; C=non-ulcerative colitis normal control patients.





**Figure 3.** ulcerative colitis progressor and nonprogressor log<sub>2</sub>-ratios at the most informative bacterial artificial chromosomes 3905, 3906, and 3907.

**Table 1**

Extent of Neoplastic Progression in 9 UC Progressors

UC Progressor Patient	Most Advanced Neoplastic Diagnosis	Dukes' Colorectal Carcinoma Stage	Incident Neoplasia vs. Prevalent Neoplasia vs. Unknown Status	Extent of Colonic Disease	Age of UC Onset	UC Disease Duration	PSC
P1	Carcinoma	Dukes' B	Incident	Rectum	61 yr	13 yr	No
P2	High-Grade Dysplasia	-	Incident	Pancolitis	7 yr	35 yr	No
P3	Carcinoma	Dukes' C	Prevalent	Pancolitis	20 yr	13 yr	No
P4	Carcinoma	Dukes' B	Unknown	Pancolitis	5 yr	35 yr	No
P5	Carcinoma	Dukes' A	Prevalent	Pancolitis	27 yr	4 yr	No
P6	High-Grade Dysplasia	-	Prevalent	Pancolitis	28 yr	29 yr	Yes
P7	Carcinoma	Dukes' B	Incident	Pancolitis	?	Long Hx	No
P8	Carcinoma	Dukes' A	Incident	Pancolitis	12 yr	21 yr	No
P9	Carcinoma	Dukes' C	Unknown	Unknown	38 yr	13 yr	Yes

yr= year; Hx = history; ?=unknown; UC=ulcerative colitis; PSC=primary sclerosing cholangitis; incident neoplasia=diagnosed while under endoscopic biopsy surveillance; prevalent neoplasia= diagnosed at initial endoscopy

**Table 2**

Chromosomal Sites with Nominally Significant Differences (All Losses) Comparing UC Progressors and Non-Progressors by Array-Comparative Genomic Hybridization ( $\alpha=0.05$  before correcting for multiple comparisons)

944	RP11-460I19	4p16	0.187	0.0015	0.001
960	RP11-101J14	4p16	0.203	0.0176	0.022
1283	RP11-94J21	5p15	0.206	0.0131	0.0133
1490	RP11-21I6	5q35	0.202	0.0158	0.035
2594	RP11-35C24	10q26	0.239	0.0168	0.0279
3476	RP11-91J9	15q24	0.19	0.0044	0.0101
3796	RP11-84J3	17q25	0.195	0.0074	0.0172
3907	RP11-315M18	18q23	0.235	1e-04	6e-04
3905	RP11-7H17	18q23	0.196	1e-04	3e-04
3906	RP11-196B3	18q23	0.238	6e-04	6e-04
4299	RP11-54I20	Xq28	0.229	0.0138	0.0279

\* BAC = bacterial artificial chromosome