High-resolution analysis of DNA copy number alterations in colorectal cancer by array-based comparative genomic hybridization

Kentaro Nakao1,*, Kshama R.Mehta2,*, Jane Fridlyand2, Dan H.Moore3, Ajay N.Jain2, Amalia Lafuente4, John W.Wiencke2, Jonathan P.Terdiman2,† and Frederic M.Waldman2,†

Email: waldman@cc.ucsf.edu

1Second Department of Surgery, Showa University School of Medicine, Tokyo, Japan, 2Comprehensive Cancer Center and 3Department of Epidemiology and Biostatistics, University of California, San Francisco, Comprehensive Cancer Center, 2340 Sutter Street, Room S436, San Francisco, CA 94143-0808, USA and 4Institut Salut Publica de Catalunya, Universitat de Barcelona, Barcelona, Spain

Array-based comparative genomic hybridization (CGH) allows for the simultaneous examination of thousands of genomic loci at 1-2 Mb resolution. Copy number alterations detected by array-based CGH can aid in the identification and localization of cancer causing genes. Here we report the results of array-based CGH in a set of 125 primary colorectal tumors hybridized onto an array consisting of 2463 bacterial artificial chromosome clones. On average, 17.3% of the entire genome was altered in our samples (8.5 ± 6.7% gained and 8.8 ± 7.3% lost). Losses involving 8p, 17p, 18p or 18q occurred in 37, 46, 49 and 60% of cases, respectively. Gains involving 8q or 20q were observed 42 and 65% of the time, respectively. A transition from loss to gain occurred on chromosome 8 between 41 and 48 Mb, with 25% of cases demonstrating a gain of 8p11 (45-53 Mb). Chromosome 8 also contained four distinct loci demonstrating high-level amplifications, centering at 44.9, 60, 92.7 and 144.7 Mb. On 20q multiple high-level amplifications were observed, centering at 32.3, 37.8, 45.4, 54.7, 59.4 and 65 Mb. Few differences in DNA copy number alterations were associated with tumor stage, location, age and sex of the patient. Microsatellite stable and unstable (MSI-H) tumors differed significantly with respect to the frequency of alterations (20 versus 5%, respectively, P < 0.01). Interestingly, MSI-H tumors were also observed to have DNA copy number alterations, most commonly involving 8q. This high-resolution analysis of DNA copy number alterations in colorectal cancer by array-based CGH allowed for the identification of many small, previously uncharacterized, genomic regions, such as on chromosomes 8 and 20. Array-based CGH was also able to identify DNA copy number changes in MSI-H tumors.

Introduction

The development and progression of colorectal cancer is a multi-step process leading to the accumulation of genomic alterations that occur over the lifetime of a tumor (1,2). Colorectal cancers can be classified by the type of genomic instability observed, with over 85% of colorectal cancers demonstrating gross chromosomal aberrations, and 10-15% of tumors demonstrating microsatellite instability (3,4). Chromosomal aberrations in tumors lead to DNA copy number alterations with associated gain or loss of genes important in tumor progression (5). Array-based comparative genomic hybridization (CGH) allows for high throughput, high-resolution genome-wide screening of DNA copy number changes in solid tumors (6-8). Copy number alterations detected by array-based CGH can be directly related to sequence information to aid in the localization, identification and validation of cancer causing genes and molecular pathways of carcinogenesis (9-11).

Array-based CGH has been applied to a number of solid tumors including breast (9), bladder (12), kidney (13), brain (14), nasopharynx (15) and esophagus (16). The technology has been shown to be useful in amplicon identification (9,11), and differential tumor diagnosis (13). Here we report on the results of array-based CGH in a set of 125 phenotypically well-characterized colorectal cancers. A high frequency of DNA copy number alterations were detected by array-based CGH involving regions of the genome documented previously to be altered in colorectal cancer (1,17,18), including major portions of chromosomes 8, 17, 18 and 20. In addition, the high-resolution analysis of DNA copy number alterations afforded by array-based CGH allowed for the identification of many small, previously uncharacterized, genomic regions, such as on chromosomes 8 and 20. Major differences in the frequency of gains, losses, amplifications and deletions were noted between tumors with and without microsatellite instability. Surprisingly, regions of chromosomal gain and loss were still seen in microsatellite unstable tumors. Some differences also were noted between tumors with different phenotypic characteristics such as tumor stage and location.

Materials and methods

Patient materials

DNA was extracted from a consecutive case-series of 125 primary colorectal cancers obtained from the University of Barcelona Hospital Clinic in Barcelona, Spain. The primary tumors were surgically dissected and immediately frozen at −80°C. DNA was extracted as described previously (19,20). Tumors have previously been characterized for microsatellite instability (19). Lymphocytic DNA from normal individuals was used as reference. Clinicopathologic data were obtained by questionnaire and review of medical records (19,20). Patients provided signed informed consent, and the study was approved by the institutional review boards at the Barcelona Hospital and the University of California San Francisco.

CGH arrays

The arrays used in the study were prepared and hybridized as described previously (12,13). Human 1.14 arrays were obtained from the University of
California, San Francisco Cancer Center Array Core. For each array, genomic target DNA was prepared by ligation-mediated PCR before being robotically spotted in triplicate onto chromium-coated glass slides. The array used in this study consisted of 2463 bacterial artificial chromosome (BAC) clones that covered the human genome at a 1.5 Mb resolution. DNA samples were hybridized to the array as described previously, with minor modifications (12,13). One microgram of tumor DNA was labeled by random priming with fluorolink cy3-dUTP, and normal reference DNA was labeled in the same fashion with cy5-dUTP (Amersham Pharmacia, Piscataway, NJ). Unincorporated fluorescent nucleotides were removed using Sephadex G-50 spin columns. Test and reference DNA were mixed with 100 µg Cot-1 DNA (Life Technologies, Gibco BRL, Gaithersburg, MD), precipitated and resuspended in 30-50 µL of a hybridization solution containing 50% formamide, 10% dextran sulfate, 2 × SCC, 4% SDS and 100 µg yeast tRNA. The hybridization solution was denatured for 10 min at 72°C before being incubated for 1 h at 37°C to allow blocking of the repetitive sequences. Hybridization was performed for 48 h in a moist chamber on a slowly rocking table, followed by washing for 15 min in 50% formamide/2 × SSC at 45°C, and 10 min in phosphate buffer at room temperature. Slides were mounted in 90% glycerol in phosphate buffer containing DAPI at a concentration of 0.3 µg/mL.

Three, 16 bit fluorescence single-color intensity images (DAPI, cy3 and cy5) were collected from each array using a charge coupled device camera (Sensys, Photometric, equipped with a Kodak KAF 1400 chip) coupled to a 1× optical system, as described previously (12,13).

Data analysis
Image data were analyzed by Spot and Sproc software as described previously (12,13). Spot exclusion criteria included the removal of unmapped clones and clones appearing in fewer than 70% of the samples. Applying these selection criteria resulted in narrowing down the clone number from 2463 clones to 2139 clones. Removal of polymorphic clones further reduced the overall clone number to 2120 clones. A list of these clones is available at http://cc.ucsf.edu/people/waldman/colon/nakao.polymorphisms.xls.

A series of 10 normal versus normal hybridizations were performed to define the low-level variation of the test to reference (T/R) log2 intensity ratio to each target clone. A clone-toclon variability in the intensity ratios was observed, but the overall coefficient of variation was <10%. The log2 ratios for each case were median centered to zero. The threshold for determining chromosome gain or loss was defined as log2 ratio >0.225 or less than −0.225. This threshold corresponds to values between 2 and 3 standard deviations from the mean. In addition, high-level amplifications were defined as a log2 ratio >0.9 and high-magnitude deletions as log2 ratio less than −0.75. These thresholds for amplification and deletion were derived from analyses of cell lines with known gene amplification and homozygous loss at defined loci. The threshold for gain or loss of an entire chromosome arm was defined as a median log2 ratio of >0.14 or less than −0.14 for all clones on the chromosome arm. The fraction of the genome gained or lost for each case was calculated as the sum of genomic distances represented by each clone.

Microsatellite instability
Tumor microsatellite instability status was determined using the BAT26 microsatellite marker using standard methods (21). Tumors were defined as having high-frequency microsatellite instability (MSI-H) if change of length of a microsatellite marker using standard methods (21). Tumors were defined as having MSI-H if change of length was observed, but the overall coefficient of variation was 0.225. This threshold corresponds to values between 2 and 3 standard deviations from the mean. In addition, high-level amplifications were defined as a log2 ratio >0.9 and high-magnitude deletions as log2 ratio less than −0.75. These thresholds for amplification and deletion were derived from analyses of cell lines with known gene amplification and homozygous loss at defined loci. The threshold for gain or loss of an entire chromosome arm was defined as a median log2 ratio of >0.14 or less than −0.14 for all clones on the chromosome arm. The fraction of the genome gained or lost for each case was calculated as the sum of genomic distances represented by each clone.

Statistics
Since multiple comparisons may result in a large number of genes appearing significant by chance alone, significance was determined by Student’s t-test, ANOVA and a MaxT test using permutation analysis to control for family-wise false positive error rates (22). Briefly, for the MaxT test, a t-statistic is computed to determine the significance of copy number changes between two groups. The group labels are then randomly permuted 1000 times and the maximum t-statistic observed between the groups is recorded. The MaxT adjusted P-value is then calculated by considering the proportion of permutations in which the maximum permutation based t-statistic exceeded the observed statistic for each clone.

Results
Array-based CGH was performed on 125 primary colorectal cancers (Figure 1). Data are available at http://cc.ucsf.edu/people/waldman/colon/nakao.data.xls. The clinical characteristics and the microsatellite instability status of the tumor samples are summarized in Table I.

Array-based CGH in colorectal cancer
By array-based CGH, an average of 360 clones were gained or lost in tumor DNA samples, comprising 17.3% of the genome (182 clones representing 8.5 ± 6.7% of the entire genome gained and 178 clones representing 8.8 ± 7.3% of the genome lost). The mean number of clones demonstrating high-level amplifications and high-magnitude deletions were 16 and 7 per case (representing 0.2 ± 0.6% of the entire genome amplified and 0.2 ± 0.8% of the genome deleted), respectively.

The majority of clones were infrequently gained or lost, with 95% of the clones being gained or lost <35% of the time. However, high-frequency gains (≥35%) were detected on 7p, 7q, 8q, 11q and 20q, and high-frequency losses were detected on 5q, 8p, 17p, 18p, 18q and 21q (Figure 2). The six most commonly altered chromosome arms were 8p loss (37%), 8q gain (42%), 17p loss (46%), 18p loss (49%), 18q loss (60%) and 20q gain (65%). The entire chromosome arm was lost in most cases involving 17p, 18p or 18q. A case-by-case analysis did not reveal smaller regions of alteration on these chromosomes. However, smaller regions of alteration were seen on chromosomes 8 and 20 as described below. Forty-nine clones demonstrated high-level amplification >5% of the time. These clones mapped to regions on 8q and 20q (see below). Only 11 clones demonstrated high-magnitude deletions >3% of the time. These clones mapped to regions on 8p (see below), 17p (15.7 Mb), 18q (21.6, 40.4, 50.2, 52.4, 86.5 and 89.2 Mb) and 21q (24.1 and 39.7 Mb).

High-resolution analysis of chromosome 8
Chromosome 8 gains or losses were very common, with 8p loss seen in 46 cases (37%) and 8q gain in 52 cases (42%). In many cases, the transition from loss to gain did not occur at the centromere (53 Mb), but occurred on 8p, between 41 and 48 Mb (Figure 3A). For clones mapping to this region of transition, the frequency of loss was 11.3%, and the frequency of gain was 17.8%. This region contains FGFR1, mapping to 44.9 Mb. For clones proximal to this region (48–53 Mb), the frequency of loss was only 2.0%, but the frequency of gain was now 25%. Case-by-case analysis did not reveal any minimally lost region on 8p or minimally gained region on 8q.

Discrete, non-contiguous regions of high-level amplification and high-magnitude deletion were identified (Figure 3B). High-magnitude deletions were relatively frequent in the region near the 8p telomere. Clone RP11-82K8 at 3.4 Mb was deleted in 5.7% of cases. High-level amplifications were common among clones at four distinct loci on chromosome 8, centered at 44.9, 60, 92.7 and 144.7 Mb (containing MYC). Representative individual cases demonstrating these regions of chromosome 8 deletion and amplification are depicted in Figure 4A–C.

High-resolution analysis of chromosome 20
Gain of 20p was observed in 33% of cases, while loss of 20p was seen in 13% of cases. Gain of 20q occurred in 65% of the cases (Figure 5A). No minimally gained region on 20q could be identified. However, there appeared to be several discrete regions across 20q that were characterized by very frequent high-level amplifications (10-15%), with intervening areas where the frequency of amplification would fall below 5% (Figure 5B). These frequent high-level amplifications were centered at 32.3, 37.8, 45.4, 54.7, 59.4 and 65 Mb, regions that contain candidate oncogenes including, ZNF217 (53.9 Mb), CYP24 (54.5 Mb) and aurora 2 kinase.
Representative individual cases demonstrating some of these DNA copy number alterations are depicted in Figure 6A-C.

**Relationship of array-based CGH with clinical phenotype**

Very minor differences in the overall frequency of DNA copy number alterations were observed across the spectrum of phenotypic characteristics such as tumor stage, location, age and sex of the patient (Table II). Right-sided tumors (proximal to the splenic flexure), and tumors from women, tended to have a smaller fraction of the genome altered than left-sided tumors, or tumors from men. In addition, no significant differences in regional alterations (chromosome arms) were observed between tumors of different clinical phenotypes.

Permutation analysis was undertaken to determine if any individual clones showed significant differences in tumor groups with differing clinical characteristics. Permutation analysis corrects for the multiple comparisons problem by...
adjusting the observed $P$ values after a random distribution of $P$ values is determined by scrambling the outcome variable for each case. After permutation analysis, no statistically significant differences were found when comparing early and late stage tumors on a clone-by-clone basis. However, statistically significant differences were found in the frequency of alterations of a small number of clones when comparing tumors from men and women, and between right- and left-sided tumors. Clone RP11-163N11 on chromosome 2 (40.4 Mb) and RP1118J23 on chromosome 19 (52.9 Mb) were significantly more likely to be gained in men than women (40.6 versus 3.5% and 12.5 versus 8.5%, respectively, MaxT = 0.001). Clone RP11-188A12 on chromosome 7 (155.7 Mb) was more likely to be gained in women than men (41.8 versus 17.0%, MaxT = 0.01), while clone RMC17P069 on chromosome 17 (21.4 Mb) was more commonly lost in women than men (29.3 versus 8.2%, MaxT = 0.006). Two clones on 6p, RP11-83B17 (123.1 Mb) and RP11-273J1 (124.0 Mb), differed significantly between left- and right-sided tumors. RP11-83B17 was gained in 9.9% of left-sided tumors, but was lost in 7.9% of right-sided cases (MaxT = 0.01). RP11-273J1 was gained in 14.5% of left-sided tumors, but was lost in 15.2% of right-sided tumors (MaxT = 0.08).

**Relationship between array-based CGH and tumor microsatellite instability status**

Only seven of the tumors demonstrated MSI-H based on analysis of the BAT26 marker. Microsatellite stable and unstable tumors differed significantly with respect to the total frequency of gains (10 versus 3%, $P < 0.01$), losses (10 versus 2%, $P < 0.01$), amplifications (0.5 versus 0%, $P < 0.01$) and deletions (0.2 versus 0.1%, $P < 0.01$). DNA copy number

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Total number of cases (% of total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage ($n = 121$)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11 (9)</td>
</tr>
<tr>
<td>2</td>
<td>37 (31)</td>
</tr>
<tr>
<td>3</td>
<td>35 (29)</td>
</tr>
<tr>
<td>4</td>
<td>38 (31)</td>
</tr>
<tr>
<td>Location ($n = 123$)</td>
<td></td>
</tr>
<tr>
<td>Righta</td>
<td>38 (31)</td>
</tr>
<tr>
<td>Left</td>
<td>85 (69)</td>
</tr>
<tr>
<td>Age ($n = 123$)</td>
<td></td>
</tr>
<tr>
<td>&lt;50 years</td>
<td>10 (8)</td>
</tr>
<tr>
<td>50-70 years</td>
<td>47 (38)</td>
</tr>
<tr>
<td>≥70 years</td>
<td>66 (54)</td>
</tr>
<tr>
<td>Sex ($n = 123$)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>59 (52)</td>
</tr>
<tr>
<td>Female</td>
<td>64 (48)</td>
</tr>
<tr>
<td>Bat26 ($n = 109$)</td>
<td></td>
</tr>
<tr>
<td>Stable</td>
<td>102 (94)</td>
</tr>
<tr>
<td>Unstable</td>
<td>7 (6)</td>
</tr>
</tbody>
</table>

aProximal to the splenic flexure.

**Fig. 2.** Overall frequency of DNA copy number alterations by array-based CGH. Frequency analysis measured as a fraction of cases gained or lost over the 2120 BAC clones. Data are presented ordered by chromosomal map position of the clones. Lower bars represent losses or deletions for all clones, and the upper bars represent gains or amplifications. The dashed lines represent the location of the centromeres.
alterations were observed among all seven MSI-H tumors, with two of seven cases having clear copy number gains of clones on 8q (Figure 7A and B). In the second case, all of chromosome 8 was gained. In the first case, gains were noted involving 8p11 and all of 8q, with the highest amplitude of gains centering at 45 Mb.

**Discussion**

Chromosomal CGH and loss of heterozygosity (LOH) analysis has been used in the past to provide for genome-wide assessment of DNA copy number alterations in colorectal tumors. However, these methods are generally of low resolution and...
labor intensive. As the resolution of array-based CGH is much higher than conventional CGH, ~1–2 Mb on these arrays compared with 10 Mb for chromosomal CGH, results from array-based CGH can greatly facilitate gene discovery. Unlike conventional CGH, array-based CGH also can provide quantitative information on the level of gain or loss. These quantitative data provide the ability to discern regions of the genome particularly likely to harbor critical cancer-related genes, such as regions with high-level amplification or high-magnitude deletion.

DNA copy number alterations were common among our 125 colorectal cancers when analyzed by array-based CGH. On average over 17% of the entire genome was altered in these tumors. However, as there was considerable variability among individual tumors, there did not appear to be a clear threshold dose of DNA that was altered in association with the development of invasive cancer. In a recent report by Hermsen and colleagues using chromosomal CGH, the number of chromosomal alterations was predictive of which adenomas had progressed to cancer and which had not (23). Their use of

Fig. 4. Continued.
chromosomal CGH, however, did not allow a quantitative assessment of the amount of genome altered in the progressed and non-progressed adenomas. Our results suggest that invasive colorectal cancer can occur over a range of genomic instability. Interestingly, in the current study, only 10% of the tumors had >36% of the genome gained or lost, and no tumors had >52% of the genome altered. The tumors in this study were not microdissected to ensure >70% tumor cell fraction, so normal cell contamination may have affected the absolute value of the upper limit of observed genomic instability, but our findings do suggest that there may be a maximal amount of genomic instability that is associated with ongoing tumor viability.

The DNA copy number alterations observed in the colorectal tumors were not random, but rather involved particular regions of the genome, most commonly involving parts or all of 8p, 8q, 17p, 18p, 18q and 20q. These results are similar to those found by our group, and others, using chromosomal CGH (23-33), or other methods, such as LOH analysis (34-36). The similarity of our findings with those of others using alternative methods, serves to validate the array-based CGH technique. Importantly, however, array-based CGH provided us with the opportunity to make a much higher resolution assessment of DNA copy number alterations simultaneously across the entire genome than is possible with chromosomal CGH or LOH studies.

DNA copy number alterations on chromosome 8 were observed in the majority of our cases, in agreement with previous reports (23,28,33). The precise loci that are targeted for loss on 8p or gain on 8q, however, remain unclear. A few studies have used LOH analysis to better define the precise region(s) of targeted loss or gain. Some of these previous studies defined minimally lost regions on 8p that may harbor tumor suppressor genes (37-40). However, our high-resolution analysis, using 152 clones on chromosome 8, could not confirm a specific region as being targeted for loss or gain. Our high-resolution analysis did, however, demonstrate several interesting features regarding genomic alterations involving chromosome 8. We demonstrated that the transition from loss to gain occurs before the centromere in the region between 41 and 48 Mb. Clones mapping to the region of 8p11 (48–53 Mb) were gained ~25% of the time, and losses were rare. Gains of this region of 8p have been reported in breast (41), bladder (42) and esophageal cancer (16), but never before in colorectal cancer. In addition, we observed high-level amplifications centering at 44.9 Mb in ~5% of our cases. This region of the genome contains the fibroblast growth factor receptor 1 gene (FGFR1). FGFR1 is a growth factor receptor mapped to overlap with three clones RP11-265K5, GS-566K20 and RP11-100B16, at 44.9 Mb. The receptor, in response to ligand binding, triggers a cascade of downstream signals influencing mitogenesis and differentiation. The role of FGFR1 in colorectal cancer has not been well studied. Jayson et al. did report the expression of FGFR1 in colonic adenoma and carcinoma cells (43).

We also demonstrated three other discrete regions of high-level amplification on 8q centering at 60, 92.7 and 144.7 Mb in our samples. MYC is partially contained by clone RP11-237F24 at 144.9 Mb (8q24.1). This clone was gained 35% of the time, and amplified with a frequency of 5.8% in our study. MYC, a downstream target of the Wnt signaling pathway, is involved in cell proliferation (44), and has been shown to be amplified or over-expressed in upwards of 50% of colorectal cancers (45-47). Other possible oncogenes that map to the regions found frequently amplified on 8q remain uncertain. Lastly, in >5% of the cases we observed a very narrow region of high-magnitude loss defined by a sub-telomeric clone at
This region is a good candidate to harbor a tumor suppressor gene. Alternatively, this observation may result from random telomeric loss in the tumors.

DNA copy number alterations involving chromosome 20 occurred in 65% of our colorectal cancer cases. Gains, and high-level amplifications of 20q, have been reported by other investigators in colorectal cancer (23,25,28,48). Gains at 20q have been associated with increased proliferative activity of colorectal cancer cells (27). High-level amplifications at 20q13.2 are significantly more common in metastatic colorectal cancer lesions than in their matched primary tumors (49–51), and reduced survival has been associated with gains.

Fig. 5. High-resolution analysis of chromosome 20. (A) Frequency of gains (triangle) and losses (filled triangle). (B) Frequency of amplifications (square) and deletions (filled square). Array-based CGH data are presented according to chromosome 20 position in Mb, with chromosome 20 banding regions and the centromere (vertical line) indicated. Genes of interest are noted.
of 20q13.2 (32). Our use of array-based CGH permitted a detailed analysis of the gains and amplifications across 20q.

We observed several common, and fairly discrete, regions of high-level amplification on 20q, centering at 32.3, 37.8, 45.4, 54.7, 59.4 and 65 Mb. The high degree of complexity of the DNA alterations on 20q in our colorectal cancers suggests that there may be multiple genes in this region that are targets for selection in colorectal cancer carcinogenesis. Gain or amplification in the region 32–38 Mb (20q11.2) has not been described previously in colorectal cancer, although it has been described in breast cancer (52). Candidate oncogenes on 20q encompassed by this and the other regions of 20q amplification found in our cases include the amplified in breast cancer gene, AIB1, located at 35 Mb (20q11.2) (53), and three candidate oncogenes on 20q13.2, ZNF217 (53.9 Mb), CYP24 (54.5 Mb) and aurora 2 kinase (56.6 Mb). These genes are known to be over-expressed in a variety of tumors, especially breast cancer (52), but their status in colorectal cancer remains

Fig. 6. Continued.
largely unstudied. ZNF217 encodes a Kruppel-like transcription factor that has been demonstrated to promote the immortalization of human mammary cells and to play a role in the suppression of apoptosis (11,54,55). Over-expression of CYP24 may abrogate growth control mediated by vitamin D or calcium (9). Aurora 2 kinase is a mitotic serine threonine kinase involved in accurate chromosomal segregation through the regulation of centrosome duplication (56). Over expression of aurora 2 kinase is associated with chromosomal instability, and has been demonstrated in over 50% of primary colorectal cancers (57).

We sought to determine if differences in the array-based CGH profile could distinguish tumors of different clinical phenotypes. We found few statistically significant differences with respect to the fraction of the genome altered, or in the particular regions of the genome changed. This result was somewhat surprising, especially as it applied to differences between tumors of different stages. Past studies have shown clearly an increase in the number of chromosome arms gained or lost with progression from adenoma or dysplasia to invasive colorectal cancer (23,25), and when comparing primary tumors with their metastases (26,29,30,33). However, some

Table II. Fraction of the genome altered by clinical phenotype

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Total number of cases</th>
<th>% Genome gained&lt;sup&gt;a&lt;/sup&gt;</th>
<th>% Genome lost</th>
<th>% Genome amplified</th>
<th>% Genome high-magnitude deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>8.9</td>
<td>6.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>8.8</td>
<td>9.2</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>8.2</td>
<td>8.6</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>4</td>
<td>38</td>
<td>8.8</td>
<td>9.8</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right&lt;sup&gt;c&lt;/sup&gt;</td>
<td>38</td>
<td>7.9</td>
<td>8.1</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Left</td>
<td>85</td>
<td>8.8</td>
<td>9.1</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50 years</td>
<td>10</td>
<td>8.0</td>
<td>9.2</td>
<td>0.4</td>
<td>0.2</td>
</tr>
<tr>
<td>50-70 years</td>
<td>47</td>
<td>9.0</td>
<td>9.3</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>&gt;70 years</td>
<td>66</td>
<td>8.1</td>
<td>8.4</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>59</td>
<td>9.2</td>
<td>9.1</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Female</td>
<td>64</td>
<td>7.7</td>
<td>8.5</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Bat26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable</td>
<td>102</td>
<td>9.4&lt;sup&gt;d&lt;/sup&gt;</td>
<td>9.6&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Unstable</td>
<td>7</td>
<td>3.3</td>
<td>2.6</td>
<td>0.0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<sup>a</sup>Fraction of genomic change was determined as the sum of the genomic distances represented by the individual BAC clones that were gained, lost, amplified or deleted.

<sup>b</sup><sup>P</sup> < 0.1 for stage 1 versus 4.

<sup>c</sup>Proximal to the splenic flexure.

<sup>d</sup><sup>P</sup> < 0.01 for Bat26 stable versus unstable.
studies (23,28,58,59), but not all (36,60,61), using chromosomal CGH and other approaches, have failed to demonstrate a significant increase in the number of genomic alterations with an increase in stage in invasive tumors (23,27–29). It is likely that chromosomal instability is an early event in colorectal carcinogenesis (62), and that the overall fraction of the genome altered is relatively constant once invasive cancer develops, at least with respect to the primary tumors. It is also possible that our result is related to an increase in normal cell contamination with later stage tumors, possibly as a result of desmoplasia. We expected to find that particular regions of genomic alteration might differ between early and late stage tumors, as particular gene alterations may be selected that confer the ability to metastasize. Our failure to find many regional, or clone specific, differences may simply be a function of the relatively small number of tumors of each stage examined, the large

![Fig. 7. High-resolution analysis of chromosome 8 in two microsatellite unstable tumors. Each tumor demonstrates copy number gains for multiple clones across 8q. In (A) a gain of a portion of 8p and all of 8q with the highest amplitude of gain centering at 45 Mb (arrow) is seen. In (B) gain involving all of chromosome 8 is seen. Array-based CGH data for chromosome 8 are presented according to chromosome 8 position in Mb, with the location of banding regions and centromere position (vertical line) indicated.](http://carcin.oxfordjournals.org/)

[Figure 7](http://carcin.oxfordjournals.org/)
number of comparisons made and the stringent statistical methods used to define significant differences.

As expected, we observed that DNA alterations detected by array-based CGH are far less common in MSI-H tumors than in microsatellite stable tumors. Previous studies have found that MSI-H tumors are generally diploid, and when studied by chromosomal CGH, they are rarely found to have DNA copy number alterations (48,63–66). However, using high-resolution array-based CGH, we were able to demonstrate that DNA copy number alterations do occur in some MSI-H tumors, especially involving chromosome 8. This finding corroborates that of Goel et al., who found LOH in 23% of MSI-H tumors (67), and suggests that MSI and chromosomal instability are not mutually exclusive mechanisms of genomic instability in all colorectal cancers. The frequency and importance of DNA copy number alterations in MSI-H colorectal cancers requires further study in a larger set of tumors.

In summary, our analysis of 125 primary colorectal cancers by array-based CGH indicates that DNA copy number alterations are common. Our use of high-resolution array-based CGH allowed us to detect the complexity of DNA alterations that occur in colorectal cancer, especially with respect to regions of high-level amplifications and high-magnitude deletions, regions that are more likely to harbor important oncogenes and tumor suppressor genes. We were able to delineate discrete regions of DNA copy number alteration on chromosomes 8 and 20 that are likely to harbor such genes. Surprisingly, the number of DNA alterations and the specific alterations varied little across tumors of differing clinical characteristics. This may be in part due to the fact that genomic instability is an early event in colorectal cancer carcinogenesis, and in part due the fact that we analyzed a relatively small number of tumors in each phenotypic class. Not surprisingly, we did find that DNA copy number alterations are far less common in microsatellite unstable than stable tumors, but our high-resolution array-based CGH was able to detect DNA copy number gains and losses in the microsatellite unstable tumors, a finding previously under appreciated when chromosomal CGH was used to analyze MSI tumors.

Acknowledgements

The authors wish to thank S.DeVries for her help in the application of CGH array technology to colorectal samples. Microarrays were provided by D.Pinkel and D.Albertson of the UCSF Cancer Center Array Core, with funding provided by NCI P30 CA 82103. This project was supported by a grant from the National Cancer Institute (CA92374).

References

by CGH and their associations with tumour DNA ploidy, genotypes and phenotypes. Br. J. Cancer, 80, 526-535.


