

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

Kentaro Nakao^{1,5}, Kshama R. Mehta^{2,5}, Jane Fridlyand², Dan H. Moore², Ajay N. Jain², Amalia Lafuente³, John W. Wiencke⁴, Jonathan P. Terdiman^{2,6}, Frederic M. Waldman^{2,6,7}

¹Second Department of Surgery, Showa University School of Medicine, Tokyo, Japan

²Comprehensive Cancer Center, and ⁴Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA

³Institut Salut Publica de Catalunya, Universitat de Barcelona, Barcelona, Spain

⁵K.N. and K.R.M. contributed equally as first-authors.

⁶J.P.T. and F.M.W. contributed equally as senior authors.

Running title: Array-based CGH in colorectal cancer

Key words: comparative genomic hybridization, colorectal cancer, chromosome 8, chromosome 20, microsatellite instability

⁷**To whom correspondence should be addressed, at:** UCSF Comprehensive Cancer Center, 2340 Sutter Street, Room S436, San Francisco, CA 94143-0808. Phone: (415) 476-3821; Fax: (415) 476-8218; Email: waldman@cc.ucsf.edu

Abbreviations: BAC, bacterial artificial chromosome; CGH, comparative genomic hybridization; Mb, megabase; MSI, microsatellite instability; MSI-H, high-frequency microsatellite instability; LOH, loss of heterozygosity.

Abstract

Array-based comparative genomic hybridization (CGH) allows for the simultaneous examination of thousands of genomic loci at 1-2 megabase 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 ($8.5\% \pm 6.7\%$ gained and $8.8\% \pm 7.3\%$ 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 to 53Mb). 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 multistep 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 previously documented 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.

Material 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 previously described [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 arrays used in this study consisted of 2463 bacterial artificial chromosome (BAC) clones that covered the human genome at a 1.5 megabase (Mb) resolution.

Each tumor DNA sample was hybridized to the array as described previously, with minor modifications [12,13]. Five hundred ng 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, 2X SSC, 4% SDS and 100 μg yeast tRNA. The hybridization solution was denatured for 10 minutes at 72°C before being incubated for 1 hour at 37°C to allow blocking of the repetitive sequences. Hybridization was performed for 48 hours in a moist chamber on a slowly rocking table, followed by washing for 10 minutes in 50% formamide/ 2X SSC at 45°C , and 10 minutes in phosphate buffer at room temperature. Slides were mounted in 90% glycerol in phosphate buffer containing DAPI at a concentration of 0.3 $\mu\text{g}/\text{ml}$.

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

Data analysis

Image data were analyzed by Spot and Sproc software as previously described [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 polymorphic clones is available at <http://cc.ucsf.edu/people/waldman/colon/nakao.polymorphisms.xls>.

A series of ten normal versus normal hybridizations were performed to define the normal variation of the test to reference (T:R) log₂ intensity ratio for each target clone. A slight clone to clone 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 < -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 < -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 < -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 mutations were detected in *BAT26* when comparing normal with tumor DNA.

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 regression analysis 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 $8.5\% \pm 6.7\%$ of the entire genome gained and 178 clones representing $8.8\% \pm 7.3\%$ of the genome lost). The mean number of clones demonstrating high level amplifications and high magnitude deletions were 16 and 7 per case (representing $0.2\% \pm 0.6\%$ of the entire genome amplified and $0.2\% \pm 0.8\%$ of the genome deleted), respectively.

The majority of clones were infrequently gained or lost, with 95% of the clones being gained or lost less than 35% of the time. However, high frequency gains ($\geq 35\%$) 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 greater than 5% of the time. These clones mapped to regions on 8q and 20q (see below). Only 11 clones demonstrated high magnitude deletions more than 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 gains 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 to 53Mb), 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 Mb, 60 Mb, 92.7 Mb, and 144.9 Mb (containing *MYC*). Representative individual cases demonstrating these regions of chromosome 8 deletion and amplification are depicted in Figures 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 Mb, 37.8Mb, 45.4 Mb, 54.7 Mb, 59.4 Mb, and 65 Mb, regions that contain candidate oncogenes including, *ZNF217* (53.9 Mb), *CYP24* (54.5 Mb), and *aurora 2 kinase* (56.6 Mb). Representative individual cases demonstrating some of these DNA copy number alterations are depicted in Figures 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 7 of the tumors demonstrated high-frequency microsatellite instability (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 alterations were observed among all 7 MSI-H tumors, with 2 of 7 cases

having clear copy number gains of clones on 8q (Figure 7 A-B). In the first case, all of chromosome 8 was gained. In the second 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. Since the resolution of array-based CGH is much higher than conventional CGH, approximately 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, since 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 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 greater than 36% of the genome gained or lost, and no tumors had greater than 52% of the genome altered. The tumors in this study were not microdissected to ensure greater than 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 Mb-53 Mb) were gained approximately 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 approximately 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 3 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 Mb, 92.7 Mb, 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 overexpressed 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 greater than 5% of the cases we observed a very narrow region of high magnitude loss defined by a sub-telomeric clone at 3.4 Mb (RP1182K8). 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 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 Mb, 37.8 Mb, 45.4 Mb, 54.7, 59.4 Mb, 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 previously described in colorectal cancer, though 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, *AIB4*, 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 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]. Overexpression 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]. Overexpression 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 clearly shown 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 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 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

1. Vogelstein, B., Fearon, E.R., Hamilton, S.R., Kern, S.E., Preisinger, A.C., Leppert, M., Nakamura, Y., White, R., Smits, A.M. and Bos, J.L. (1988) Genetic alterations during colorectal-tumor development. *N Engl J Med*, **319**, 525-32.
2. Fearon, E.R. and Vogelstein, B. (1990) A genetic model for colorectal tumorigenesis. *Cell*, **61**, 759-767.
3. Kinzler, K.W. and Vogelstein, B. (1996) Lessons from hereditary colorectal cancer. *Cell*, **87**, 159-170.
4. Lengauer, C., Kinzler, K.W. and Vogelstein, B. (1998) Genetic instabilities in human cancers. *Nature*, **396**, 643-9.
5. Knuutila, S., Bjorkqvist, A.M., Autio, K., Tarkkanen, M., Wolf, M., Monni, O., Szymanska, J., Larramendy, M.L., Tapper, J., Pere, H., El-Rifai, W., Hemmer, S., Wasenius, V.M., Vidgren, V. and Zhu, Y. (1998) DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. *Am J Pathol*, **152**, 1107-23.
6. Pollack, J.R., Perou, C.M., Alizadeh, A.A., Eisen, M.B., Pergamenschikov, A., Williams, C.F., Jeffrey, S.S., Botstein, D. and Brown, P.O. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nature Genetics*, **23**, 41-6.
7. Pinkel, D., Se Graves, R., Sudar, D., Clark, S., Poole, I., Kowbel, D., Collins, C., Kuo, W.L., Chen, C., Zhai, Y., Dairkee, S.H., Ljung, B.M., Gray, J.W. and Albertson, D.G. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nature Genetics*, **20**, 207-11.
8. Snijders, A.M., Nowak, N., Se Graves, R., Blackwood, S., Brown, N., Conroy, J., Hamilton, G., Hindle, A.K., Huey, B., Kimura, K., Law, S., Myambo, K., Palmer, J., Ylstra, B., Yue, J.P., Gray, J.W., Jain, A.N., Pinkel, D. and Albertson, D.G. (2001) Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet*, **29**, 263-4.
9. Albertson, D.G., Ylstra, B., Se Graves, R., Collins, C., Dairkee, S.H., Kowbel, D., Kuo, W.L., Gray, J.W. and Pinkel, D. (2000) Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene. *Nat Genet*, **25**, 144-6.
10. Cheung, V.G., Nowak, N., Jang, W., Kirsch, I.R., Zhao, S., Chen, X.N., Furey, T.S., Kim, U.J., Kuo, W.L., Olivier, M., Conroy, J., Kasprzyk, A., Massa, H., Yonescu, R., Sait, S., Thoreen, C., Snijders, A., Lemyre, E., Bailey, J.A., Bruzel, A., Burrill, W.D., Clegg, S.M., Collins, S., Dhami, P., Friedman, C., Han, C.S., Herrick, S., Lee, J., Ligon, A.H., Lowry, S., Morley, M., Narasimhan, S., Osoegawa, K., Peng, Z., Plajzer-Frick, I., Quade, B.J., Scott, D., Sirotkin, K., Thorpe, A.A., Gray, J.W., Hudson, J., Pinkel, D., Ried, T., Rowen, L., Shen-Ong, G.L., Strausberg, R.L., Birney, E., Callen, D.F., Cheng, J.F., Cox, D.R.,

- Doggett, N.A., Carter, N.P., Eichler, E.E., Haussler, D., Korenberg, J.R., Morton, C.C., Albertson, D., Schuler, G., de Jong, P.J. and Trask, B.J. (2001) Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature*, **409**, 953-8.
11. Collins, C., Volik, S., Kowbel, D., Ginzinger, D., Ylstra, B., Cloutier, T., Hawkins, T., Predki, P., Martin, C., Wernick, M., Kuo, W.L., Alberts, A. and Gray, J.W. (2001) Comprehensive genome sequence analysis of a breast cancer amplicon. *Genome Res*, **11**, 1034-42.
 12. Veltman, J.A., Fridlyand, J., Pejavar, S., Olshen, A.B., Korkola, J.E., DeVries, S., Carroll, P., Kuo, W.L., Pinkel, D., Albertson, D., Cordon-Cardo, C., Jain, A.N. and Waldman, F.M. (2003) Array-based Comparative Genomic Hybridization for Genome-Wide Screening of DNA Copy Number in Bladder Tumors. *Cancer Res*, **63**, 2872-80.
 13. Wilhelm, M., Veltman, J.A., Olshen, A.B., Jain, A.N., Moore, D.H., Presti, J.C., Jr., Kovacs, G. and Waldman, F.M. (2002) Array-based comparative genomic hybridization for the differential diagnosis of renal cell cancer. *Cancer Res*, **62**, 957-60.
 14. Hui, A.B., Lo, K.W., Yin, X.L., Poon, W.S. and Ng, H.K. (2001) Detection of multiple gene amplifications in glioblastoma multiforme using array-based comparative genomic hybridization. *Lab Invest*, **81**, 717-23.
 15. Hui, A.B., Lo, K.W., Leung, S.F., Teo, P., Fung, M.K., To, K.F., Wong, N., Choi, P.H., Lee, J.C. and Huang, D.P. (1999) Detection of recurrent chromosomal gains and losses in primary nasopharyngeal carcinoma by comparative genomic hybridisation. *Int J Cancer*, **82**, 498-503.
 16. Ishizuka, T., Tanabe, C., Sakamoto, H., Aoyagi, K., Maekawa, M., Matsukura, N., Tokunaga, A., Tajiri, T., Yoshida, T., Terada, M. and Sasaki, H. (2002) Gene amplification profiling of esophageal squamous cell carcinomas by DNA array CGH. *Biochem Biophys Res Commun*, **296**, 152-5.
 17. Vogelstein, B., Fearon, E.R., Kern, S.E., Hamilton, S.R., Preisinger, A.C., Nakamura, Y. and White, R. (1989) Allelotype of colorectal carcinomas. *Science*, **244**, 207-11.
 18. Boland, C.R., Sato, J., Appelman, H.D., Bresalier, R.S. and Feinberg, A.P. (1995) Microallelotyping defines the sequence and tempo of allelic losses at tumour suppressor gene loci during colorectal cancer progression. *Nature Medicine*, **1**, 902-9.
 19. Wiencke, J.K., Zheng, S., Lafuente, A., Lafuente, M.J., Grudzen, C., Wrensch, M.R., Miike, R., Ballesta, A. and Trias, M. (1999) Aberrant methylation of p16INK4a in anatomic and gender-specific subtypes of sporadic colorectal cancer. *Cancer Epidemiol Biomarkers Prev*, **8**, 501-6.
 20. Zheng, S., Chen, P., McMillan, A., Lafuente, A., Lafuente, M.J., Ballesta, A., Trias, M. and Wiencke, J.K. (2000) Correlations of partial and extensive methylation at the p14(ARF)

locus with reduced mRNA expression in colorectal cancer cell lines and clinicopathological features in primary tumors. *Carcinogenesis*, **21**, 2057-64.

21. Hoang, J.M., Cottu, P.H., Thuille, B., Salmon, R.J., Thomas, G. and Hamelin, R. (1997) BAT-26, an indicator of the replication error phenotype in colorectal cancers and cell lines. *Cancer Res*, **57**, 300-3.
22. Olshen, A.B. and Jain, A.N. (2002) Deriving quantitative conclusions from microarray expression data. *Bioinformatics*, **18**, 961-70.
23. Hermsen, M., Postma, C., Baak, J., Weiss, M., Rapallo, A., Sciutto, A., Roemen, G., Arends, J.W., Williams, R., Giaretti, W., De Goeij, A. and Meijer, G. (2002) Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. *Gastroenterology*, **123**, 1109-19.
24. Ried, T., Knutzen, R., Steinbeck, R., Blegen, H., Schrock, E., Heselmeyer, K., du Manoir, S. and Auer, G. (1996) Comparative genomic hybridization reveals a specific pattern of chromosomal gains and losses during the genesis of colorectal tumors. *Genes Chromosomes Cancer*, **15**, 234-45.
25. Meijer, G.A., Hermsen, M.A., Baak, J.P., van Diest, P.J., Meuwissen, S.G., Belien, J.A., Hoovers, J.M., Joenje, H., Snijders, P.J. and Walboomers, J.M. (1998) Progression from colorectal adenoma to carcinoma is associated with non-random chromosomal gains as detected by comparative genomic hybridisation. *J Clin Pathol*, **51**, 901-9.
26. Al-Mulla, F., Keith, W.N., Pickford, I.R., Going, J.J. and Birnie, G.D. (1999) Comparative genomic hybridization analysis of primary colorectal carcinomas and their synchronous metastases. *Genes Chromosomes Cancer*, **24**, 306-14.
27. De Angelis, P.M., Clausen, O.P., Schjolberg, A. and Stokke, T. (1999) Chromosomal gains and losses in primary colorectal carcinomas detected by CGH and their associations with tumour DNA ploidy, genotypes and phenotypes. *Br J Cancer*, **80**, 526-35.
28. Aust, D.E., Willenbacher, R.F., Terdiman, J.P., Ferrell, L.D., Chang, C.G., Moore, D.H., 2nd, Molinaro-Clark, A., Baretton, G.B., Loehrs, U. and Waldman, F.M. (2000) Chromosomal alterations in ulcerative colitis-related and sporadic colorectal cancers by comparative genomic hybridization. *Human Pathology*, **31**, 109-14.
29. Nakao, K., Shibusawa, M., Ishihara, A., Yoshizawa, H., Tsunoda, A., Kusano, M., Kurose, A., Makita, T. and Sasaki, K. (2001) Genetic changes in colorectal carcinoma tumors with liver metastases analyzed by comparative genomic hybridization and DNA ploidy. *Cancer*, **91**, 721-6.
30. Aragane, H., Sakakura, C., Nakanishi, M., Yasuoka, R., Fujita, Y., Taniguchi, H., Hagiwara, A., Yamaguchi, T., Abe, T., Inazawa, J. and Yamagishi, H. (2001)

Chromosomal aberrations in colorectal cancers and liver metastases analyzed by comparative genomic hybridization. *Int J Cancer*, **94**, 623-9.

31. Rooney, P.H., Boonsong, A., McKay, J.A., Marsh, S., Stevenson, D.A., Murray, G.I., Curran, S., Haites, N.E., Cassidy, J. and McLeod, H.L. (2001) Colorectal cancer genomics: evidence for multiple genotypes which influence survival. *Br J Cancer*, **85**, 1492-8.
32. De Angelis, P.M., Stokke, T., Beigi, M., Mjaland, O. and Clausen, O.P. (2001) Prognostic significance of recurrent chromosomal aberrations detected by comparative genomic hybridization in sporadic colorectal cancer. *Int J Colorectal Dis*, **16**, 38-45.
33. Knosel, T., Petersen, S., Schwabe, H., Schluns, K., Stein, U., Schlag, P.M., Dietel, M. and Petersen, I. (2002) Incidence of chromosomal imbalances in advanced colorectal carcinomas and their metastases. *Virchows Arch*, **440**, 187-94.
34. Tomlinson, I., Ilyas, M., Johnson, V., Davies, A., Clark, G., Talbot, I. and Bodmer, W. (1998) A comparison of the genetic pathways involved in the pathogenesis of three types of colorectal cancer. *J Pathol*, **184**, 148-52.
35. Zhou, W., Galizia, G., Lieto, E., Goodman, S.N., Romans, K.E., Kinzler, K.W., Vogelstein, B., Choti, M.A. and Montgomery, E.A. (2001) Counting alleles reveals a connection between chromosome 18q loss and vascular invasion. *Nat Biotechnol*, **19**, 78-81.
36. Choi, S.W., Lee, K.J., Bae, Y.A., Min, K.O., Kwon, M.S., Kim, K.M. and Rhyu, M.G. (2002) Genetic classification of colorectal cancer based on chromosomal loss and microsatellite instability predicts survival. *Clin Cancer Res*, **8**, 2311-22.
37. Fujiwara, Y., Emi, M., Ohata, H., Kato, Y., Nakajima, T., Mori, T. and Nakamura, Y. (1993) Evidence for the presence of two tumor suppressor genes on chromosome 8p for colorectal carcinoma. *Cancer Res*, **53**, 1172-4.
38. Farrington, S.M., Cunningham, C., Boyle, S.M., Wyllie, A.H. and Dunlop, M.G. (1996) Detailed physical and deletion mapping of 8p with isolation of YAC clones from tumour suppressor loci involved in colorectal cancer. *Oncogene*, **12**, 1803-8.
39. Chughtai, S.A., Crundwell, M.C., Cruickshank, N.R., Affie, E., Armstrong, S., Knowles, M.A., Takle, L.A., Kuo, M., Khan, N., Phillips, S.M., Neoptolemos, J.P. and Morton, D.G. (1999) Two novel regions of interstitial deletion on chromosome 8p in colorectal cancer. *Oncogene*, **18**, 657-65.
40. Lerebours, F., Olschwang, S., Thuille, B., Schmitz, A., Fouchet, P., Laurent-Puig, P., Boman, F., Flejou, J.F., Monges, G., Paraf, F., Bedossa, P., Sabourin, J.C., Salmon, R.J., Parc, R. and Thomas, G. (1999) Deletion mapping of the tumor suppressor locus involved in colorectal cancer on chromosome band 8p21. *Genes Chromosomes Cancer*, **25**, 147-53.

41. Dib, A., Adelaide, J., Chaffanet, M., Imbert, A., Le Paslier, D., Jacquemier, J., Gaudray, P., Theillet, C., Birnbaum, D. and Pebusque, M.J. (1995) Characterization of the region of the short arm of chromosome 8 amplified in breast carcinoma. *Oncogene*, **10**, 995-1001.
42. Simon, R., Richter, J., Wagner, U., Fijan, A., Bruderer, J., Schmid, U., Ackermann, D., Maurer, R., Alund, G., Knonagel, H., Rist, M., Wilber, K., Anabitarte, M., Hering, F., Hardmeier, T., Schonenberger, A., Flury, R., Jager, P., Fehr, J.L., Schraml, P., Moch, H., Mihatsch, M.J., Gasser, T. and Sauter, G. (2001) High-throughput tissue microarray analysis of 3p25 (RAF1) and 8p12 (FGFR1) copy number alterations in urinary bladder cancer. *Cancer Res*, **61**, 4514-9.
43. Jayson, G.C., Vives, C., Paraskeva, C., Schofield, K., Coutts, J., Fleetwood, A. and Gallagher, J.T. (1999) Coordinated modulation of the fibroblast growth factor dual receptor mechanism during transformation from human colon adenoma to carcinoma. *Int J Cancer*, **82**, 298-304.
44. He, T.C., Sparks, A.B., Rago, C., Hermeking, H., Zawel, L., da Costa, L.T., Morin, P.J., Vogelstein, B. and Kinzler, K.W. (1998) Identification of c-MYC as a target of the APC pathway [see comments]. *Science*, **281**, 1509-12.
45. Smith, D.R. and Goh, H.S. (1996) Overexpression of the C-Myc Proto-Oncogene in Colorectal Carcinoma Is Associated With a Reduced Mortality That Is Abrogated By Point Mutation of the P53 Tumor Suppressor Gene. *Clinical Cancer Research*, **2**, 1049-1053.
46. Obara, K., Yokoyama, M., Asano, G. and Tanaka, S. (2001) Evaluation of myc and chromosome 8 copy number in colorectal cancer using interphase cytogenetics. *Int J Oncol*, **18**, 233-9.
47. Kakisako, K., Miyahara, M., Uchino, S., Adachi, Y. and Kitano, S. (1998) Prognostic significance of c-myc mRNA expression assessed by semi-quantitative RT-PCR in patients with colorectal cancer. *Oncol Rep*, **5**, 441-5.
48. Schlegel, J., Stumm, G., Scherthan, H., Bocker, T., Zirngibl, H., Ruschoff, J. and Hofstadter, F. (1995) Comparative genomic in situ hybridization of colon carcinomas with replication error. *Cancer Res*, **55**, 6002-5.
49. Nanashima, A., Yamaguchi, H., Yasutake, T., Sawai, T., Kusano, H., Tagawa, Y., Nakagoe, T. and Ayabe, H. (1997) Gain of chromosome 20 is a frequent aberration in liver metastasis of colorectal cancers. *Dig Dis Sci*, **42**, 1388-93.
50. Korn, W.M., Yatsutake, T., Kuo, W.-L., Warren, R., Collins, C., Tomita, M., Gray, J. and Waldman, F.M. (1999 (in press)) Chromosome arm 20q gains and other genomic alterations in colorectal cancer metastatic to liver, as analyzed by comparative genomic hybridization and fluorescence in situ hybridization. *Genes Chromosomes Cancer*.

51. Hidaka, S., Yasutake, T., Takeshita, H., Kondo, M., Tsuji, T., Nanashima, A., Sawai, T., Yamaguchi, H., Nakagoe, T., Ayabe, H. and Tagawa, Y. (2000) Differences in 20q13.2 copy number between colorectal cancers with and without liver metastasis. *Clin Cancer Res*, **6**, 2712-7.
52. Hodgson, J.G., Chin, K., Collins, C. and Gray, J.W. (2003) Genome amplification of chromosome 20 in breast cancer. *Breast Cancer Res Treat*, **78**, 337-45.
53. Guan, X.Y., Meltzer, P.S., Dalton, W.S. and Trent, J.M. (1994) Identification of cryptic sites of DNA sequence amplification in human breast cancer by chromosome microdissection. *Nat Genet*, **8**, 155-61.
54. Collins, C., Rommens, J.M., Kowbel, D., Godfrey, T., Tanner, M., Hwang, S.I., Polikoff, D., Nonet, G., Cochran, J., Myambo, K., Jay, K.E., Froula, J., Cloutier, T., Kuo, W.L., Yaswen, P., Dairkee, S., Giovanola, J., Hutchinson, G.B., Isola, J., Kallioniemi, O.P., Palazzolo, M., Martin, C., Ericsson, C., Pinkel, D., Gray, J.W. and et al. (1998) Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 8703-8.
55. Nonet, G.H., Stampfer, M.R., Chin, K., Gray, J.W., Collins, C.C. and Yaswen, P. (2001) The ZNF217 gene amplified in breast cancers promotes immortalization of human mammary epithelial cells. *Cancer Res*, **61**, 1250-4.
56. Dutertre, S., Descamps, S. and Prigent, C. (2002) On the role of aurora-A in centrosome function. *Oncogene*, **21**, 6175-83.
57. Bischoff, J.R., Anderson, L., Zhu, Y., Mossie, K., Ng, L., Souza, B., Schryver, B., Flanagan, P., Clairvoyant, F., Ginther, C., Chan, C.S., Novotny, M., Slamon, D.J. and Plowman, G.D. (1998) A homologue of *Drosophila* aurora kinase is oncogenic and amplified in human colorectal cancers. *Embo Journal*, **17**, 3052-65.
58. Albe, X., Vassilakos, P., Helfer-Guarnori, K., Givel, J.C., de Quay, N., Suardet, L., Eliason, J.F., and Odartchenko, N. (1990) Independent prognostic value of ploidy in colorectal cancer. A prospective study using image cytometry. *Cancer*, **66**, 1168-75.
59. Weber, J.C., Schneider, A., Rohr, S., Nakano, H., Bachellier, P., Mechine, A., Hamel, G., Kanor, M., Chenard, M.P., Gaub, M.P., Oudet, P., and Jaeck, D. (2001) Analysis of allelic imbalance in patients with colorectal cancer according to stage and presence of synchronous liver metastases. *Ann Surg*, **234**, 795-802.
60. Giaretti, W., Venesio, T., Sciutto, A., Prevosto, C., Geido, E., Risio, M. (2003) Near-diploid and near-triploid human sporadic colorectal adenocarcinoma differ for KRAS2 and TP53 mutational status. *Genes Chromosomes Cancer*, **37**, 207-13.

61. De Angelis, P.M., Clausen, O.P.F., Schjolberg A., Stokke T. (1999) Chromosomal gains and losses in primary colorectal carcinomas detected by CGH and their associations with tumor DNA ploidy, genotypes and phenotypes. *Br J Cancer*, **80**, 526-535.
62. Shih I.M., Zhou W., Goodman, S.N., Lengauer, C., Kinzler, K.W., Vogelstein, B. (2001) Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis. *Cancer Research*, **61**, 818-22.
63. Thibodeau, S.N., French, A.J., Cunningham, J.M., Tester, D., Burgart, L.J., Roche, P.C., McDonnell, S.K., Schaid, D.J., Vockley, C.W., Michels, V.V., Farr, G.H., Jr. and O'Connell, M.J. (1998) Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. *Cancer Research*, **58**, 1713-8.
64. Georgiades, I.B., Curtis, L.J., Morris, R.M., Bird, C.C. and Wyllie, A.H. (1999) Heterogeneity studies identify a subset of sporadic colorectal cancers without evidence for chromosomal or microsatellite instability. *Oncogene*, **18**, 7933-40.
65. Curtis, L.J., Georgiades, I.B., White, S., Bird, C.C., Harrison, D.J. and Wyllie, A.H. (2000) Specific patterns of chromosomal abnormalities are associated with RER status in sporadic colorectal cancer. *J Pathol*, **192**, 440-5.
66. Gafa, R., Maestri, I., Matteuzzi, M., Santini, A., Ferretti, S., Cavazzini, L. and Lanza, G. (2000) Sporadic colorectal adenocarcinomas with high-frequency microsatellite instability. *Cancer*, **89**, 2025-37.
67. Goel, A., Arnold, C.N., Niedzwiecki, D., Chang, D.K., Ricciardiello, L., Carethers, J.M., Dowell, J.M., Wasserman, L., Compton, C., Mayer, R.J., Bertagnolli, M.M. and Boland, C.R. (2003) Characterization of sporadic colon cancer by patterns of genomic instability. *Cancer Res*, **63**, 1608-14.

Table I. Clinical Characteristics

Phenotype		Total Number of Cases (% of Total)
Stage (n=121)	1	11 (9%)
	2	37 (31%)
	3	35 (29%)
	4	38 (31%)
Location (n=123)	Right ^a	38 (31%)
	Left	85 (69%)
Age (n=123)	< 50 years	10 (8%)
	50 – 70 years	47 (38%)
	≥ 70 years	66 (54%)
Sex (n=123)	Male	59 (52%)
	Female	64 (48%)
Bat26 (n=109)	Stable	102 (94%)
	Unstable	7 (6%)

^a Proximal to the splenic flexure

Table II. Fraction of the genome altered by clinical phenotype

Phenotype		Total Number of Cases	Genome Gained ^a	Genome Lost	Genome Amplified	Genome Deleted
Stage	1	11	8.9%	6.1% ^b	0.1%	0.0%
	2	37	8.8%	9.2%	0.3%	0.3%
	3	35	8.2%	8.6%	0.1%	0.2%
	4	38	8.8%	9.8%	0.3%	0.2%
Location	Right ^c	38	7.9%	8.1%	0.2%	0.2%
	Left	85	8.8%	9.1%	0.3%	0.2%
Age	< 50 years	10	8.0%	9.2%	0.4%	0.2%
	50-70 years	47	9.0%	9.3%	0.3%	0.2%
	> 70 years	66	8.1%	8.4%	0.2%	0.3%
Sex	Male	59	9.2%	9.1%	0.3%	0.3%
	Female	64	7.7%	8.5%	0.2%	0.2%
Bat26	Stable	102	9.4% ^d	9.6% ^d	0.3%	0.2%
	Unstable	7	3.3%	2.6%	0.0%	0.1%

^aFraction 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.

^b $p < 0.1$ for stage 1 vs. 4

^cProximal to the splenic flexure

^d $p < 0.01$ for Bat26 stable vs. unstable

FIGURE LEGENDS

Fig. 1. Representative array-based CGH profiles from normal colorectal DNA (A) and tumor DNA (B). Profiles show copy number changes relative to normal, sex matched DNA ordered by chromosome. Clones demonstrating chromosomal gains (\log_2 ratio > 0.225), losses (\log_2 ratio < -0.225), high level amplifications (\log_2 ratio > 0.9) and high magnitude deletions (\log_2 ratio < -0.75) are indicated. Note that the array-based CGH profile of normal colorectal DNA does not show any copy number alterations. Also note the high level amplifications on 8q and 20q in the tumor sample (circled spots).

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.

Fig. 3. High resolution analysis of chromosome 8. (A) Frequency of gains (Δ) and losses (\blacktriangle); (B) Frequency of amplifications (\square) and deletions (\blacksquare). Array-based CGH data for chromosome 8 is presented according to chromosome 8 BAC position in Mb, with the location of banding regions and selected genes indicated. Centromere position is indicated by the vertical line.

Fig. 4. Three representative cases showing DNA copy number alterations involving chromosome 8. (A) Arrow indicates transition from gain to loss at 48.1 Mb. (B) Arrow indicates high magnitude deletion near the 8p telomere at 3.4 Mb. (C) Arrow indicates discrete high level amplification on 8q24 at 144.7 Mb. Array-based CGH data for chromosome 8 is presented according to chromosome 8 BAC position in Mb, with the location of banding regions and centromere (vertical line) indicated.

Fig. 5. High resolution analysis of chromosome 20. (A) Frequency of gains (Δ) and losses (\blacktriangle). (B) Frequency of amplifications (\square) and deletions (\blacksquare). Array-based CGH data is presented according to chromosome 20 BAC position in Mb, with chromosome 20 banding regions and the centromere (vertical line) indicated. Genes of interest are noted.

Fig. 6. Three representative cases showing DNA alterations involving chromosome 20. In (A) both 20p and 20q are gained, in (B) there is a gain of all of 20q, and in (C) note the complexity of amplifications across 20q, with 3 discrete regions of amplification centering at 32 Mb, 45 Mb and 53 Mb (arrows). Array-based CGH data is presented according to chromosome 20 BAC position in Mb, with chromosome 20 banding regions and the centromere (vertical line) indicated.

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 is presented according to chromosome 8 BAC position in Mb, with the location of banding regions and centromere position (vertical line) indicated.

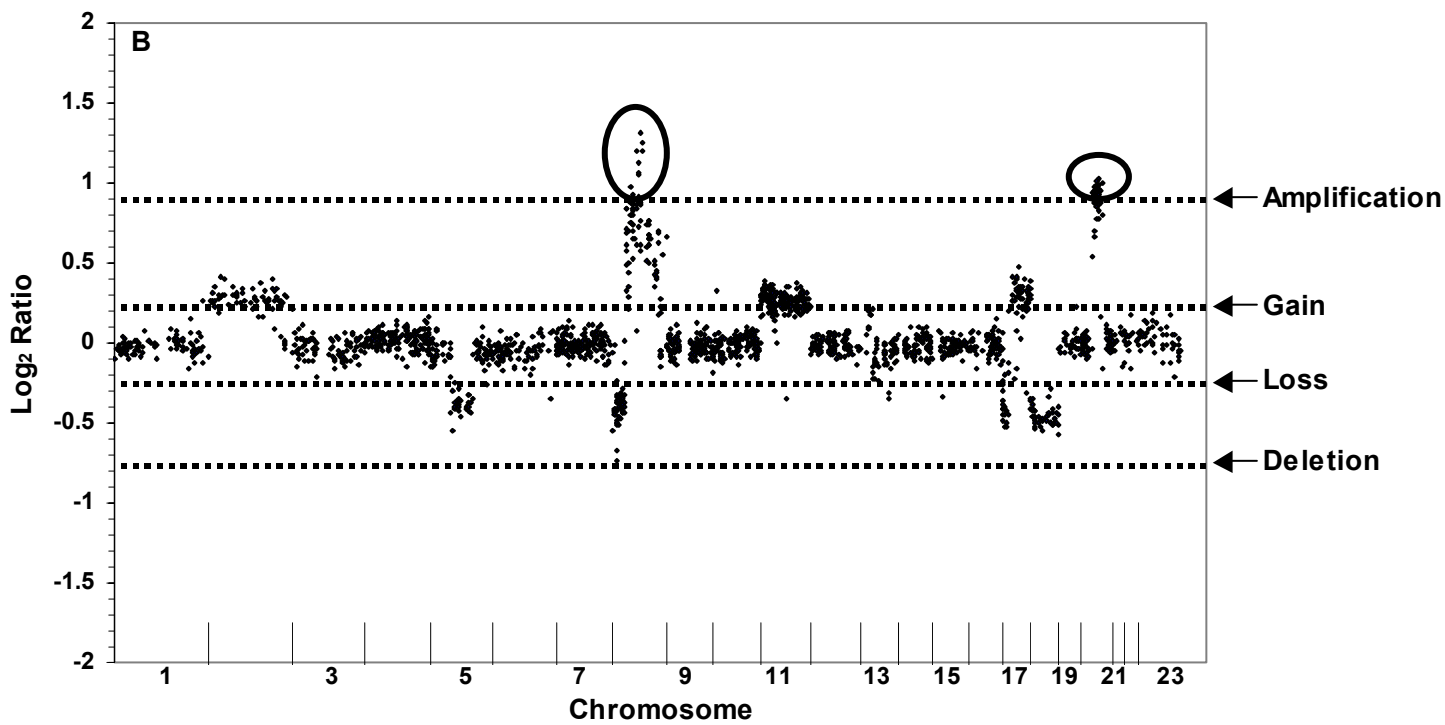
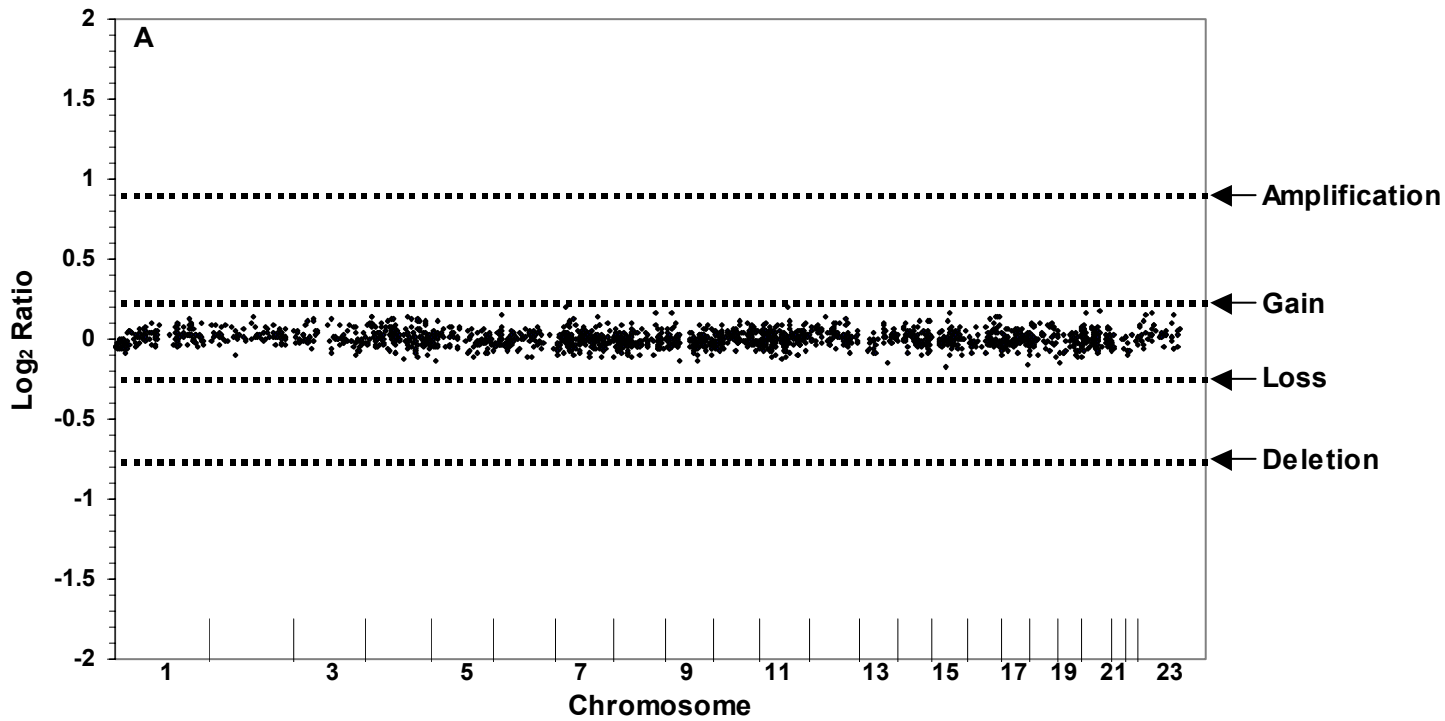


Figure 1

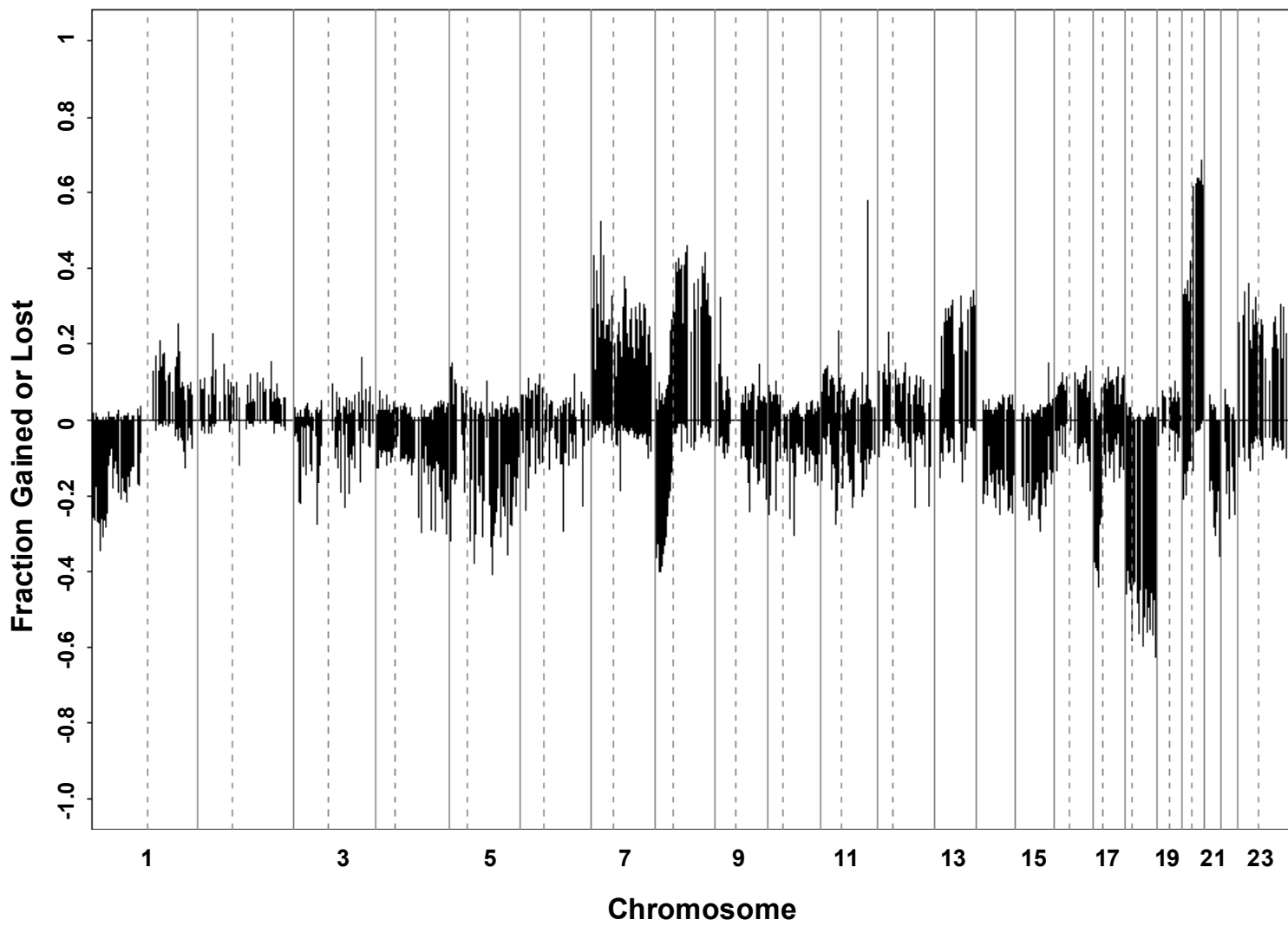


Figure 2

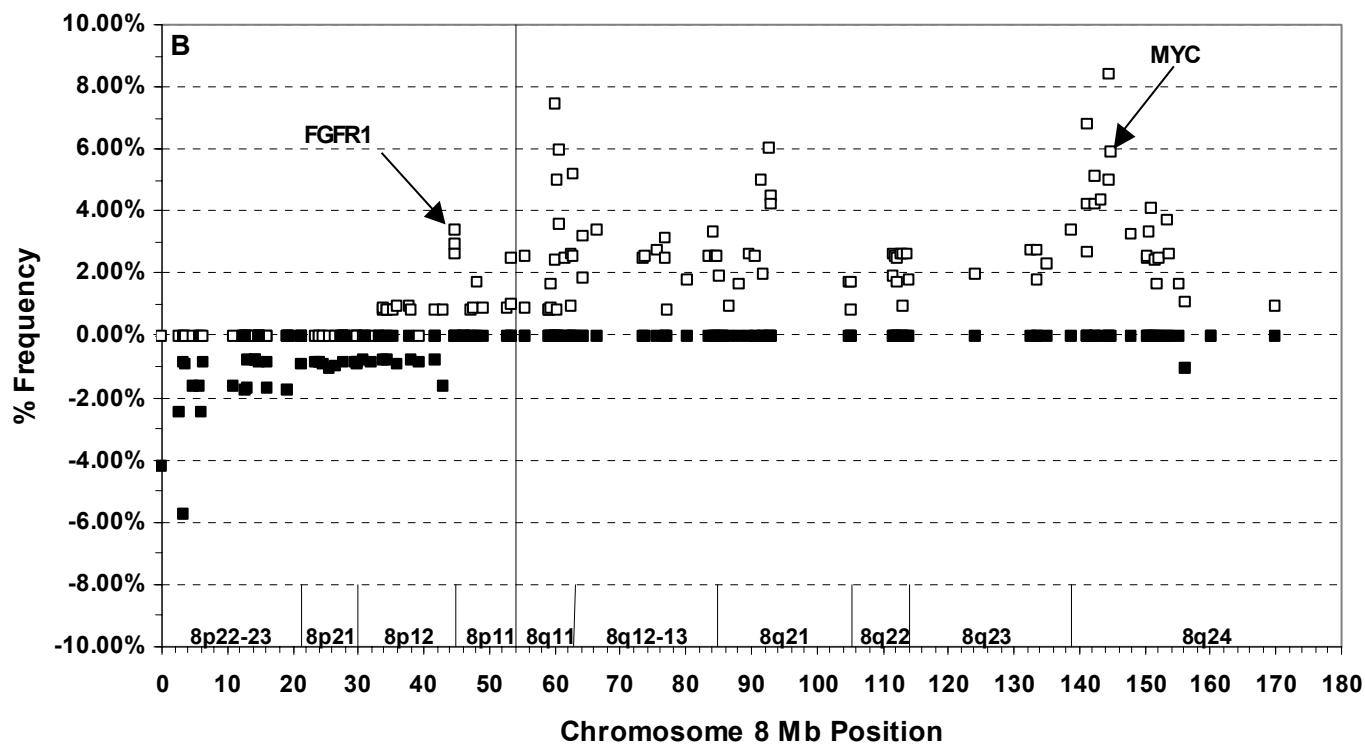
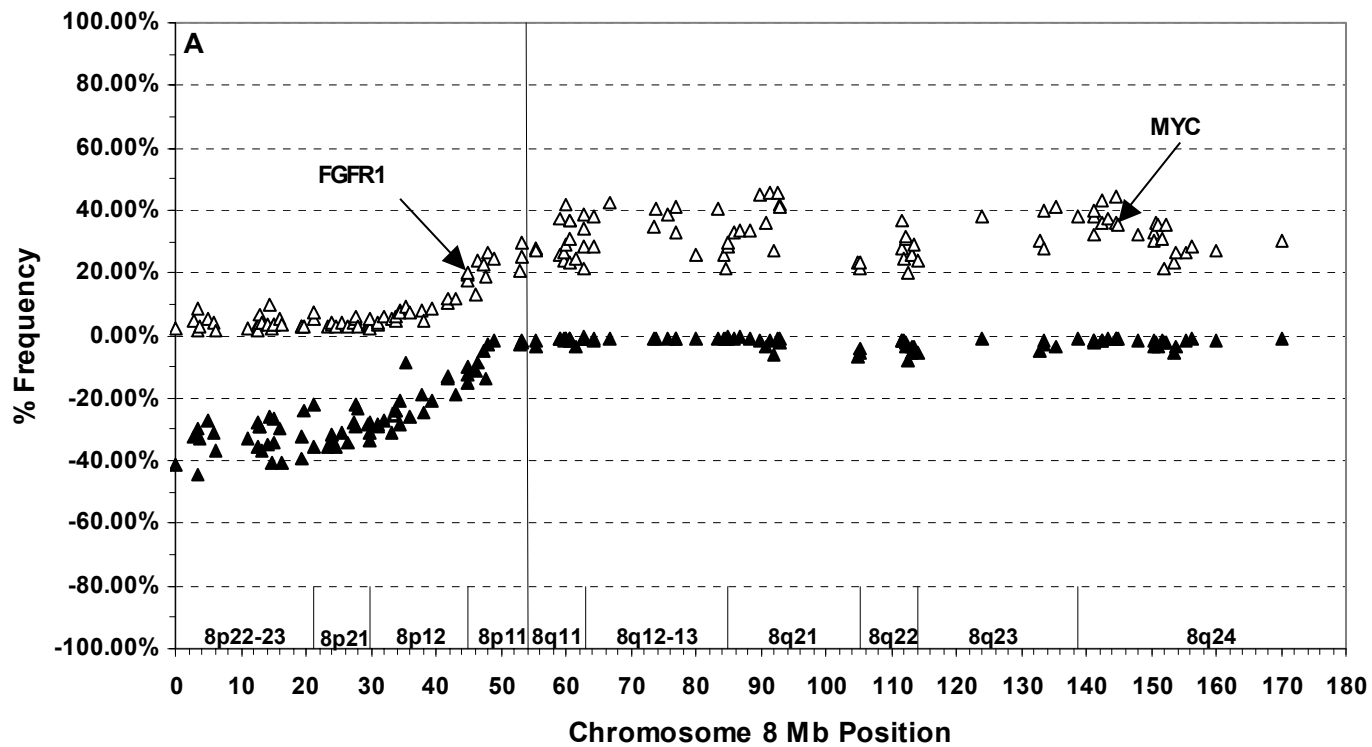


Figure 3

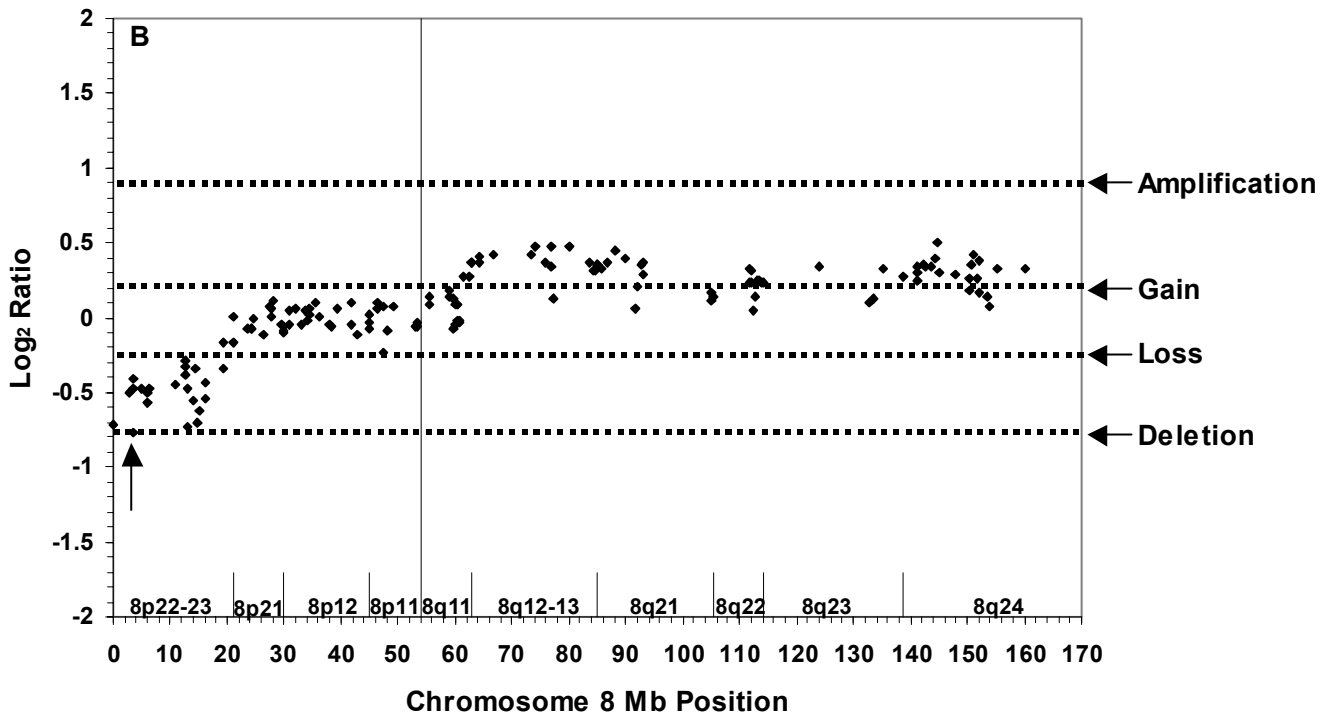
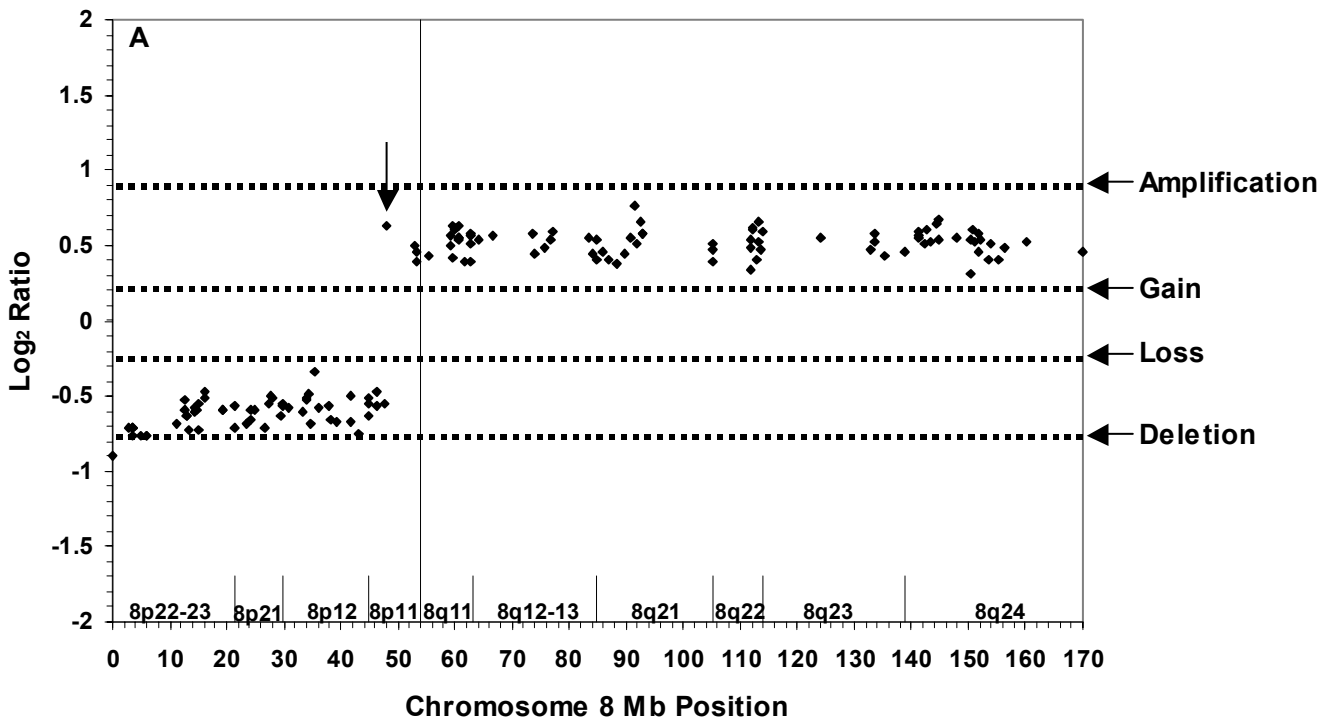


Figure 4

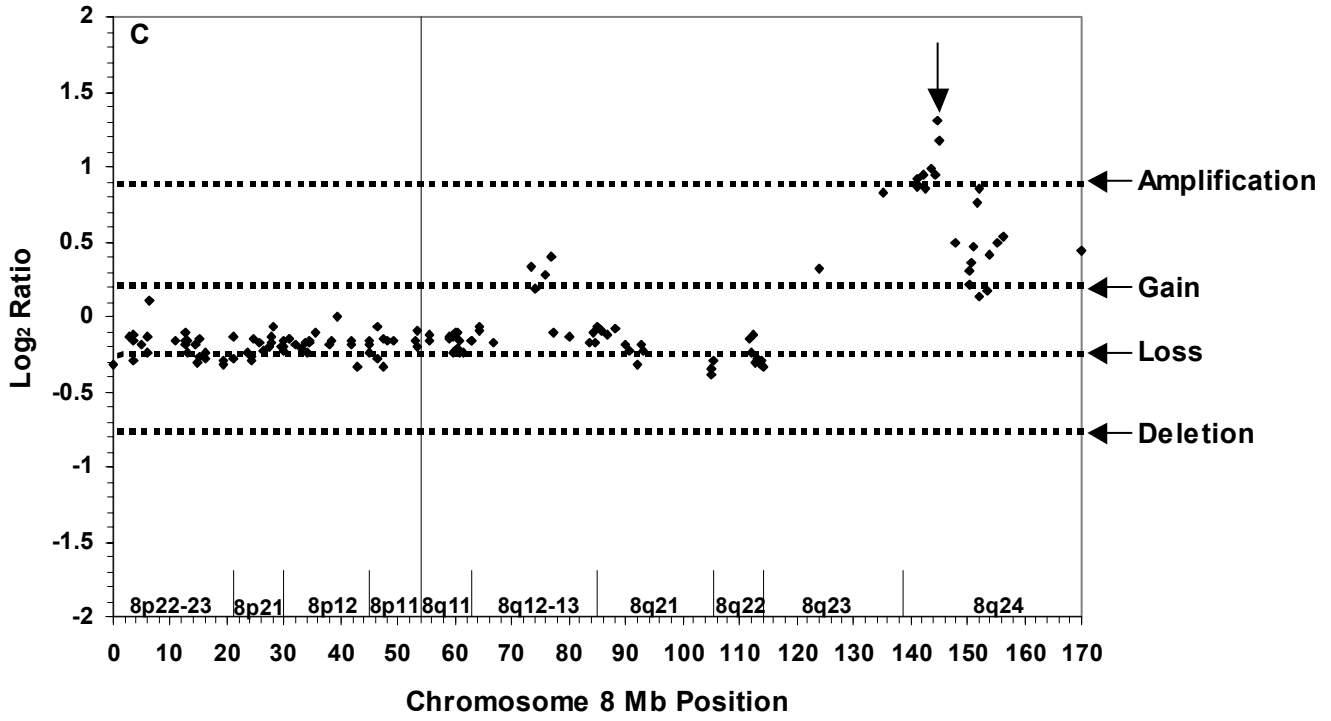


Figure 4

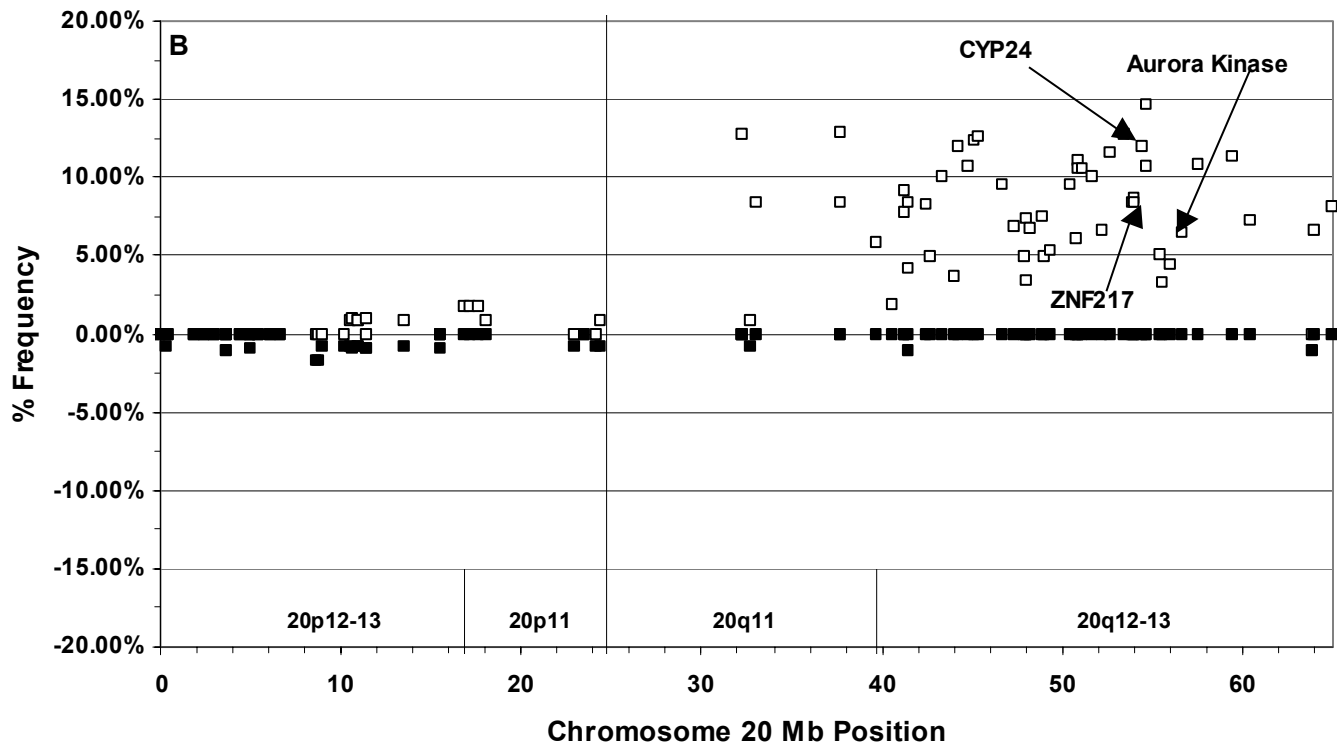
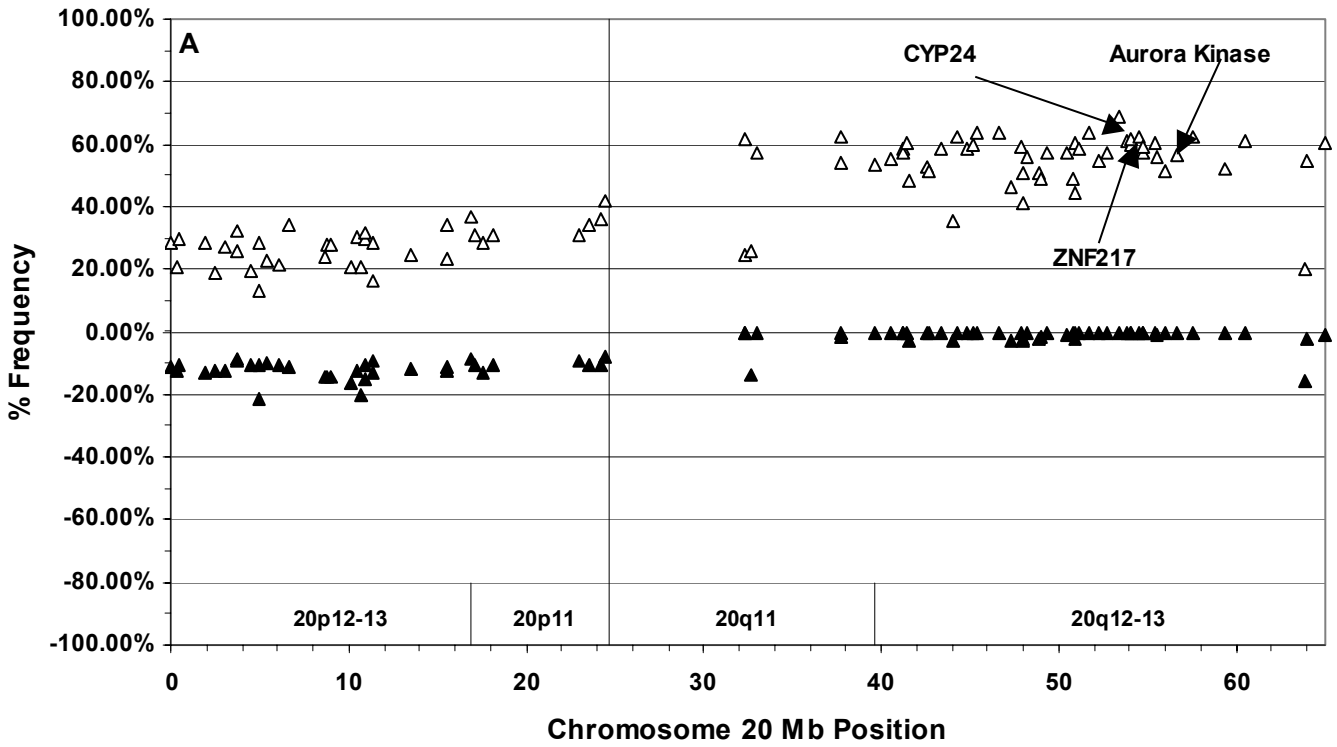


Figure 5

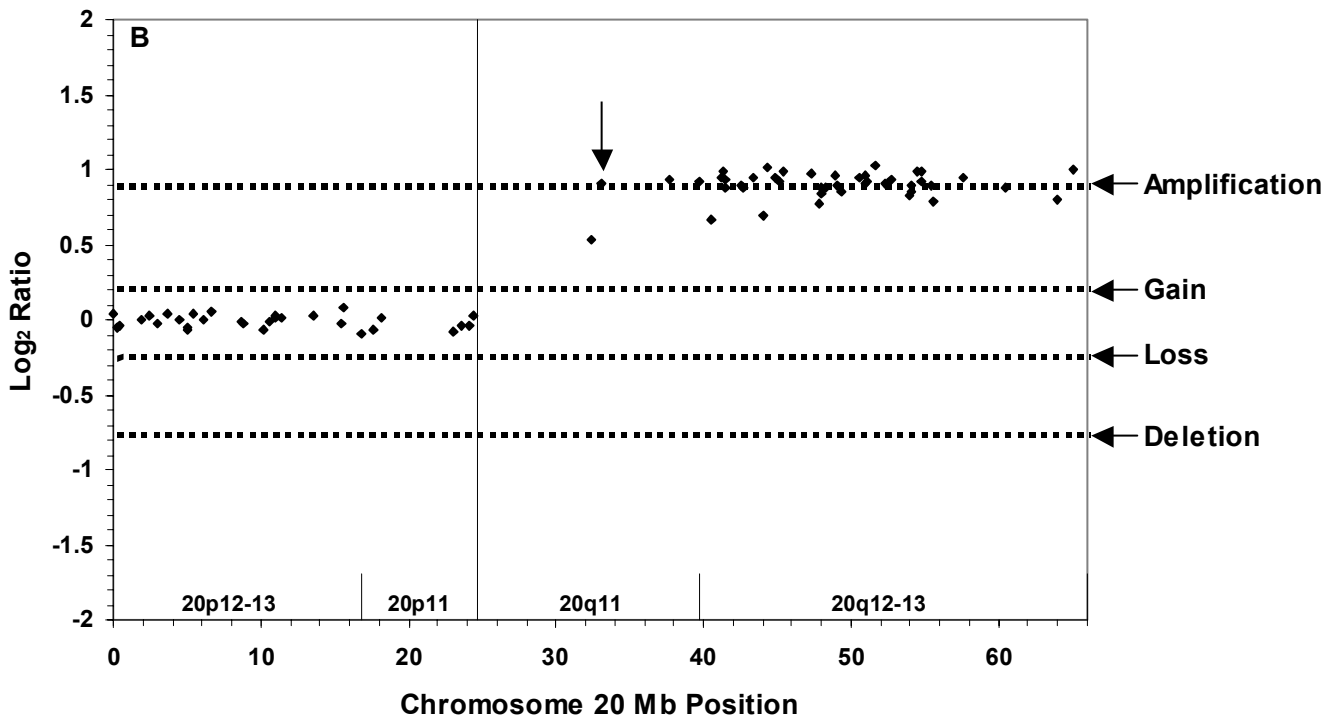
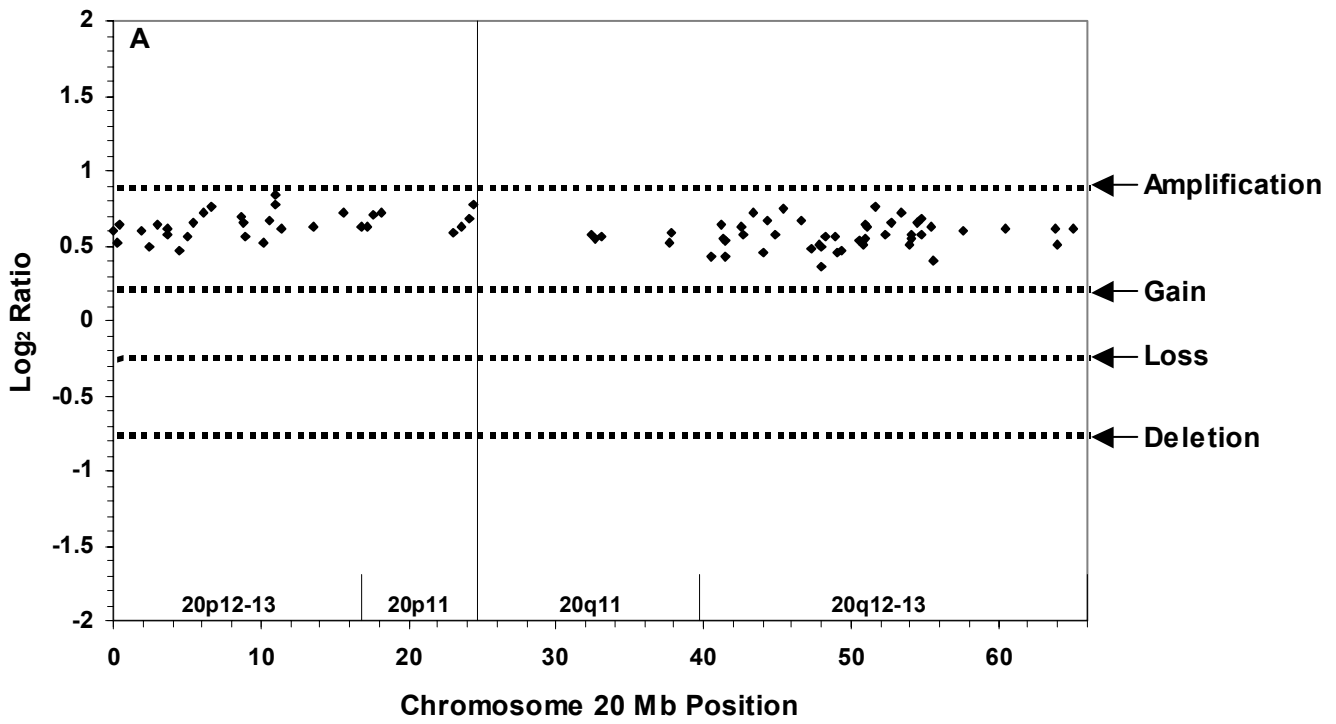


Figure 6

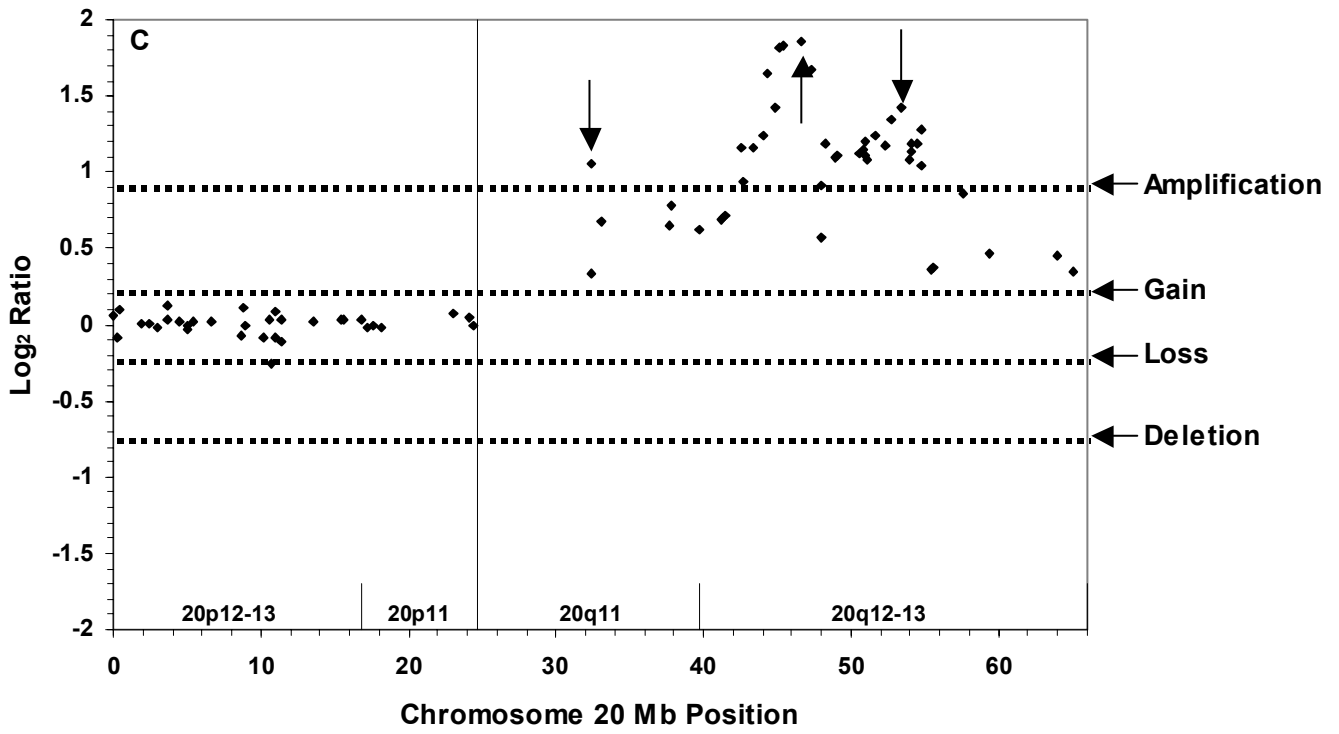


Figure 6

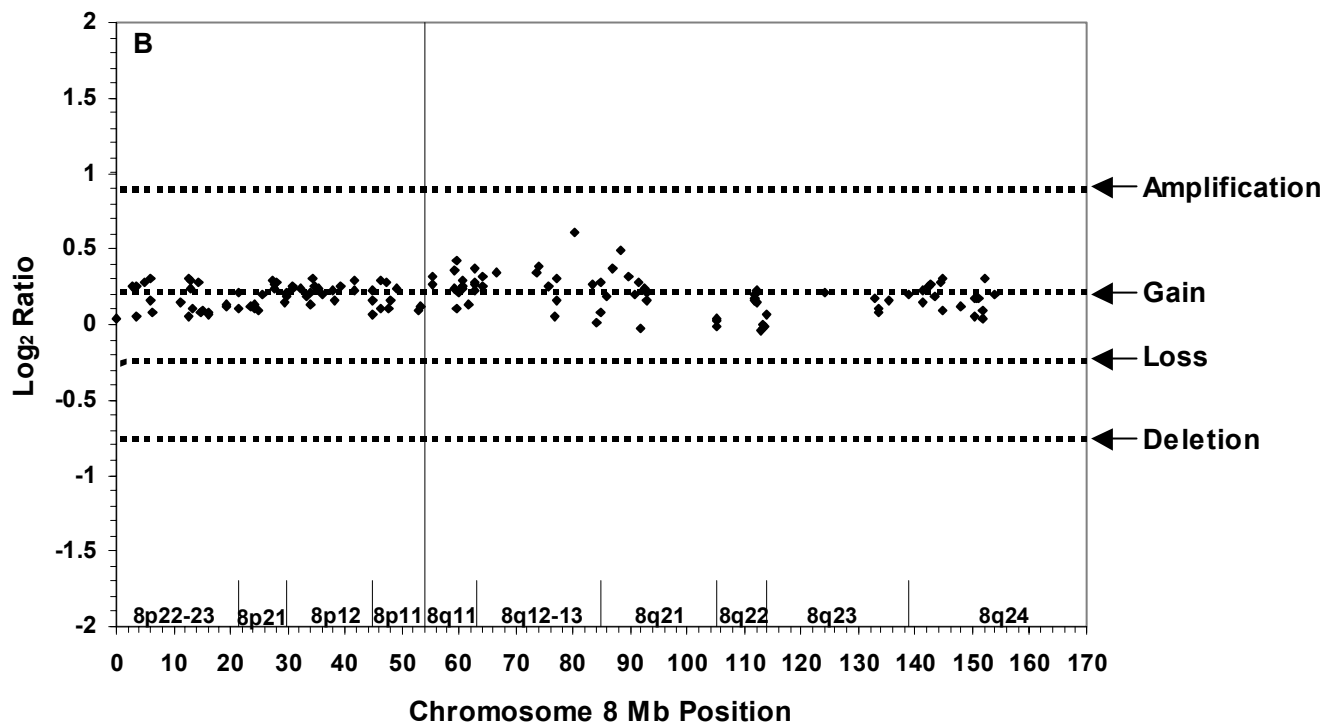
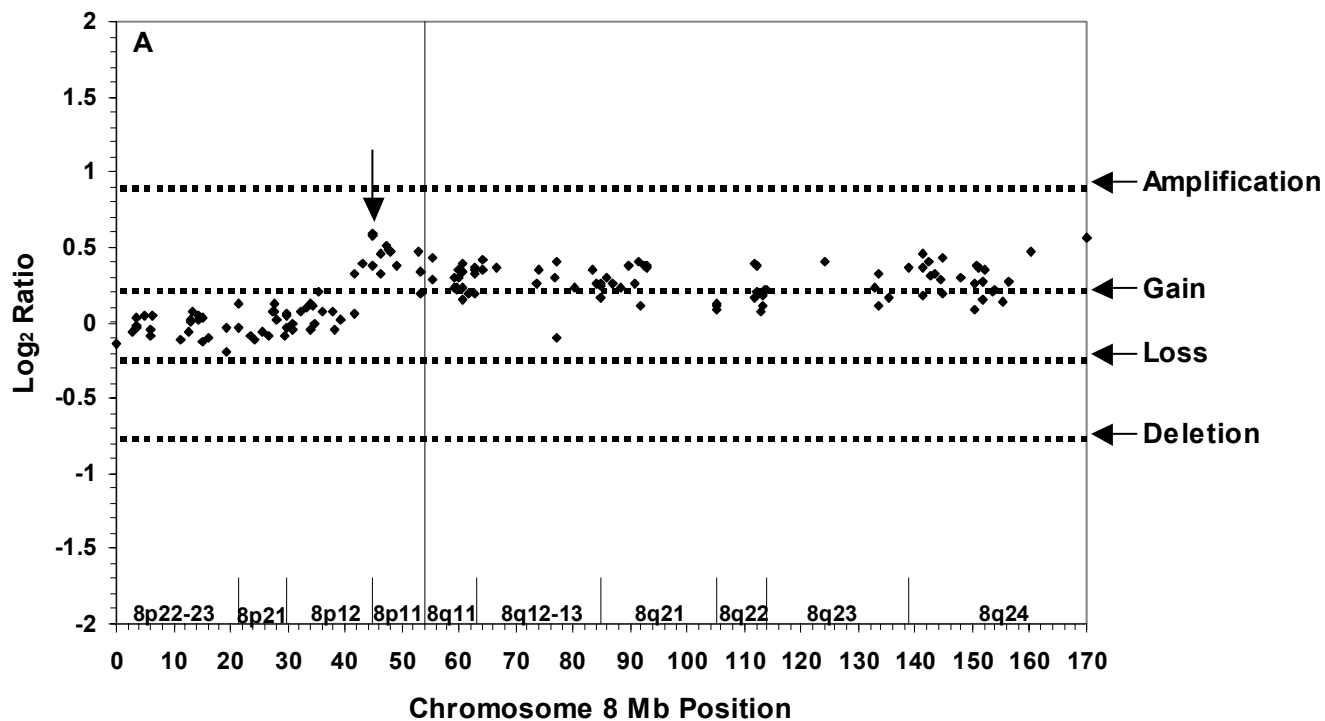


Figure 7