Chromosomal instability and copy number alterations in Barrett's esophagus and esophageal adenocarcinoma

Thomas G. Paulson[1,2,10], Carlo C. Maley[4], Xiaohong Li[2], Hongzhe Li[5], Carissa A. Sanchez[1,2], Dennis L. Chao[3], Robert D. Odze[6], Thomas L. Vaughan[2,7], Patricia L. Blount[1,2,9], Brian J. Reid[1,2,8,9]

[1]Divisions of Human Biology, [2]Public Health Sciences, and the [3]Vaccine and Infectious Disease Institute, Fred Hutchinson Cancer Research Center, Seattle, WA; [4]Molecular and Cellular Oncogenesis Program, Systems Biology Division, The Wistar Institute, Philadelphia, PA and Genomics and Computational Biology Graduate Program and Cellular and Molecular Biology Graduate Program, University of Pennsylvania, Philadelphia, PA; [5]Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, PA; [6]Department of Pathology, Brigham and Women's Hospital, Boston, MA; [7]Departments of Epidemiology, [8]Genome Sciences and [9]Medicine, University of Washington, Seattle, WA; [10]Corresponding author Brian J. Reid, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave North, C1-157, Seattle, WA, 98109. Supported by NIH K07CA089147, PO1CA91955 and the Ryan Hill Foundation (TGP).

Running Title: Copy number alterations in Barrett's esophagus

Key words: copy number alteration, Barrett's esophagus, esophageal adenocarcinoma, array CGH, premalignant, aneuploidy, chromosomal instability

Statement of Clinical Relevance

Barrett's esophagus (BE) is the only known precursor to esophageal adenocarcinoma (EA), but the vast majority of patients with BE will die of unrelated causes. Identification of biomarkers that discriminate between patients at low vs. high risk of progressing to cancer is necessary to improve patient outcomes. Here, we report an array comparative genomic hybridization (CGH) analysis of copy number alterations in a cohort of 98 patients with either premalignant BE or EA. In addition to determining the frequency and locations of deletions and amplifications occurring *before* the development of cancer, genome wide analysis of copy number alterations can identify DNA content aneuploid populations, as well as patients at risk for progression to DNA content abnormalities or EA. Array CGH provides a single platform for validation of chromosomal instability as a biomarker for cancer risk assessment that can be further evaluated in larger studies.

ABSTRACT

Purpose: Chromosomal instability, as assessed by many techniques, including DNA content aneuploidy, LOH, and comparative genomic hybridization, has consistently been reported to be common in cancer and rare in normal tissues. Recently, a panel of chromosome instability biomarkers, including LOH and DNA content, has been reported to identify patients at high and low risk of progression from Barrett's esophagus (BE) to esophageal adenocarcinoma (EA), but required multiple platforms for implementation. Although chromosomal instability involving amplifications and deletions of chromosome regions have been observed in nearly all cancers, copy number alterations (CNAs) in

premalignant tissues have not been well characterized or evaluated in cohort studies as biomarkers of cancer risk. **Experimental Design:** We examined CNAs in 98 patients having either BE or EA using BAC array CGH to characterize CNAs at different stages of progression ranging from early BE to advanced EA. **Results:** CNAs were rare in early stages (<HGD) but were progressively more frequent and larger in later stages (HGD and EA), including high level amplifications. The number of CNAs correlated highly with DNA content aneuploidy. Patients whose biopsies contained CNAs involving more than 70 Mbp were at increased risk of progression to DNA content abnormalities or EA (HR=4.9, 95% CI 1.6-14.8, p=0.0047), and the risk increased as more of the genome was affected. **Conclusions:** Genome wide analysis of CNAs provides a common platform for evaluation of chromosome instability for cancer risk assessment as well as identification of common regions of alteration that can be further studied for biomarker discovery.

INTRODUCTION

Barrett's esophagus (BE) is a premalignant condition in which the squamous epithelium that normally lines the esophagus is replaced with an intestinal metaplasia as a result of chronic gastroesophageal reflux disease (GERD). Patients with BE have at least a 15-fold increased risk for development of esophageal adenocarcinoma (EA)[1], a cancer that has increased in incidence by more than 600% over the past 30 years[2]. Treatment options for EA are limited, and the majority of patients who develop EA present initially with advanced disease, with 5 year survival rates of 13.7%[3]. Patients with BE are typically placed in surveillance programs for the early detection of cancer, but the rate of progression from BE to EA is estimated to be only 0.7% per year[4], and the vast majority

of BE patients will neither develop nor die from EA[5]. Thus, there is a strong clinical need for biomarkers that can discriminate between those who are unlikely to progress to cancer, who should be reassured and removed from frequent surveillance because of their low risk, and those at higher risk, who need frequent surveillance or intervention to prevent cancer.

Chromosomal instability involving DNA copy number alterations (CNAs) are frequently observed in many types of cancer, including those of the pancreas, lung, colon, breast and prostate, among others[6]. CNAs have been used as biomarkers for cancer prognosis in multiple studies[7, 8], but there are few longitudinal studies of CNAs as predictors of progression to cancer. Most studies analyzing CNAs that occur during neoplastic progression *in vivo* examine primarily cancer samples. CNAs in patients with EA have been examined primarily by traditional comparative genomic hybridization (CGH)[9-20]. Traditional CGH studies of EA have typically reported widespread alterations throughout the genome, but with low resolution with respect to specific chromosomal regions being affected. Recently, Nancarrow et al reported a study of EA using SNP arrays, confirming widespread and extensive chromosomal alterations in advanced EAs[21].

The utility of CNAs as biomarkers of risk assessment for progression to EA at earlier stages of neoplastic progression in BE, however, has not been well studied. Two groups have examined a small number of premalignant BE samples using traditional CGH. Croft et al, found copy number gains on multiple chromosomes in at least 40% of 15 high-grade dysplasias (HGD)[22], while Riegman et al, found frequent gains and losses in ten

HGD and nine low-grade dysplasia (LGD) samples, with no alterations observed in ten metaplasias[23]. These studies were limited by the lack of resolution of traditional CGH and the fact that the Riegman study only examined premalignant BE in specimens in which cancer had already arisen. In a more recent small study of six selected patients whose CDKN2A and TP53 status was known, it was demonstrated that changes in chromosomal instability (LOH and CNAs) could be detected over time, but, like the other studies, these patients and samples were highly selected and were not representative of the spectrum of BE in patients in general[24]. While these studies focused upon discovery of specific chromosomal alterations, well designed biomarker validations studies will be required to bring chromosome instability biomarkers to the clinic[25] A recent study evaluated a panel of tumor suppressor genes and DNA content biomarkers, including CDKN2A (LOH, methylation, mutation), TP53 (LOH, mutation), tetraploidy and aneuploidy[26]. Only the chromosome instability biomarkers, 9p LOH, 17p LOH, tetraploidy and aneuploidy, provided independent cancer risk assessment in multivariate analysis. However, this panel required a combination of platforms, including short tandem repeat polymorphisms for LOH and DNA content flow cytometry, which would be difficult to implement clinically.

Here we report for the first time evaluation of genome–wide chromosome instability analysis of copy number alterations using BAC array CGH in 174 samples from a cohort of 98 patients with diagnoses ranging from BE negative for dysplasia to advanced EA, a population representative of the range of BE stages of neoplastic progression and a sample size that provides statistical power to quantify early and relatively rare CNA

events. BAC array CGH allows genome wide analysis of copy number alterations and much more precise location of gains and deletions than traditional CGH[27]. DNA content flow cytometric data and patient characteristics were also available for each of the samples allowing us to validate array CGH as a measure of aneuploidy, a previously validated biomarker of progression from BE to EA[28]. We further investigated array CGH as a common platform to assess chromosomal instability in a prospective biomarker validation study. This study extends previous discovery research from many sources into a translational research cohort study[25] demonstrating that genome wide assessment of copy number identifies BE patients with an increased risk for progression.

METHODS

Study Subjects and Tissue Acquisition

The Seattle Barrett's Esophagus Study was approved by the Human Subjects Division of the University of Washington in 1983 and renewed annually thereafter with reciprocity from the Fred Hutchinson Cancer Research Center (FHCRC) Institutional Review Board from 1993 to 2001. Since 2001, the study has been approved by the FHCRC IRB with reciprocity from the University of Washington Human Subjects Division. The 72 non-cancer participants in this study (Table 1) had their baseline endoscopy performed between 1995 and 1999 and were followed for a period of six to 140 months. Patients were categorized on the basis of maximal histology at baseline and were grouped into three categories: less than high grade dysplasia (<HGD), which includes diagnoses of metaplasia without dysplasia, indefinite for dysplasia and low grade dysplasia; high grade dysplasia (HGD) and esophageal adenocarcinoma (EA). These categories were chosen

based upon observer variation studies, which show best reproducibility when diagnoses were divided between HGD/EA and low-grade/indefinite/metaplasia[29, 30], and upon prospective studies that show risk of progression to EA is markedly greater for HGD than for lower grades[31, 32]. EA samples came from esophagectomy specimens. The distribution of patients in this study by gender, age, BE segment length, percentage of patients progressing to EA during follow-up and histologic diagnosis is similar to that of the overall Seattle Barrett's Esophagus Cohort, with the exception of a lower percentage of patients with 17p LOH. This lower representation is due to the amount of DNA required for analysis by BAC array, which precluded the use of some samples. Forty-two of the 98 patients (43%) and 40 of the 83 (48%) non-EA patients in this study had more than one sample available for analysis (23 patients had two samples, nine had three, and five patients each had four and five samples). Different biopsies from six of the patients with EA, and different biopsies from separate endoscopies from two patients with HGD and eight of the <HGD patients were examined for genetic alterations using SNP arrays in a study published previously[33]; however, the current study was designed independently.

Endoscopic biopsy protocols used in the Seattle Barrett's Esophagus Study have been published previously[26]. Briefly, four quadrant biopsies for histology were taken every 1 cm (for patients with high-grade dysplasia and DNA content tetraploidy or aneuploidy) or every 2 cm (for patients without high-grade dysplasia or DNA content tetraploidy or aneuploidy) at intervals ranging from every 6 months to 3 years, as described previously. Additional biopsies at levels adjacent to those used for histologic evaluation were taken every 2 cm for molecular analyses; a subset of these was used in this study. Although the

biopsies used for CGH were not evaluated for histology, they came from within a region of the columnar-lined esophagus identified by an expert Barrett's endoscopist (PLB) that was histologically verified as Barrett's esophagus by an expert GI pathologist (RDO)[34]. All biopsies examined in this study were taken from either the baseline endoscopy or from a surgical resection. Endoscopic biopsies were placed into cryovials with media with 10% DMSO (dimethyl sulfoxide) held on wet ice until frozen and stored at -70°C.

Ki67/DNA Content Multiparameter Flow Cytometry and Sorting

Frozen endoscopic biopsies were prepared for flow cytometry as described previously[26]. The suspension of unfixed nuclei from each biopsy was distributed into separate tubes with approximately 10% for DNA content flow cytometric analysis and 90% for multiparameter Ki67/DNA content cell sorting. The DAPI (10 µg/ml, Accurate Chemical, Westbury, NY) saturated nuclei for single parameter DNA content flow cytometry were never centrifuged and were syringed using a 25 gauge needle immediately before acquisition on the flow cytometer. DNA content analysis was performed using MultiCycle software (Phoenix Flow Systems, San Diego, CA) with a peak vs. area gate to exclude doublets and with "sliced nucleus" background correction. The remaining nuclei were incubated with DAPI and either directly conjugated Ki67–RPE (phycoerythrin) or isotype control–RPE (DAKO R0840, Carpinteria, CA) and cell sorted to purify the proliferating BE epithelial cells from non-proliferating G_0 cells into cell cycle fractions including G1, 4N (G2/tetraploid), or aneuploid populations as previously described[26].

Array characteristics. Characteristics and construction details of the BAC arrays used in this study have been described previously[35]. The BAC arrays consist of 4342 BAC clones with median spacing 402 kb spotted in duplicate, with 99% of map locations verified by FISH. The identity and locations of individual BACs in the array can be found at the CHORI BAC/PAC resources website (FISH Mapped Clones V1.3 Download).

BAC array preparation. Probe labeling and hybridization conditions have been described previously[35]. Ten nanograms of digested genomic DNA were used as input into labeling reactions for each biopsy sample and labeled with Cy5. A single male reference DNA (Promega, Madison, WI) was used as a normal control for all samples and labeled with Cy3. The use of a single normal control raises the possibility that constitutive copy number variations may be misinterpreted as somatic genetic events[36]. We have examined the most frequent alterations described in Table 4a and have noted those that overlap with regions found to have CNV in at least one analyzed population at a frequency greater than 10% (Database of Genomic Variants[37]).

Preliminary BAC array data processing. Arrays were scanned with a GenePix 4000A scanner (Axon Instruments, Union City, CA) and data were processed using GenePix 3.0 image analysis software. Log_2 ratio of sample fluorescence to control (Cy5/Cy3) for each spot on the array was determined and all ratios were normalized and corrected for intensity-based location adjustment using a block-level loess algorithm[38]. The average log_2 ratio for the duplicate spots was determined for each BAC on the array: in cases where one of the duplicates failed, the log_2 ratio was calculated from the remaining spot.

Any BACs for which the duplicates differed by more than 20% were classified as no data. Any arrays having more than 20% bad spots were not included in the analysis.

BAC array data analysis. Statistical methods were applied to identify CNAs in the background of potentially noisy $\log_2$ ratios generated in the array CGH experiments. The wavelet method described by Hsu, et al[39], was used to denoise the BAC array data, help identify BACs with CNAs and the breakpoints of each CNA event. The wavelets method is a spatially adaptive nonparametric method that can accommodate the abrupt changes in copy numbers and different sizes of aberrations. It has been demonstrated that the wavelets-based data denoising yields greater power in the downstream statistical analyses and generates more comparable $\log_2$ ratios across samples than raw data. The predicted $\log_2$ ratios after wavelets denoising were then used to determine the calls for each BAC as a) copy number loss, b) copy number gain, c) no change or d) no data. The $\log_2$ ratio for each BAC in a sample was plotted along its position on each chromosome and the regions that were called gain and loss identified. Contiguous regions of loss, defined as a continuous region of BACs all having the same call of copy number gain or loss, were called gain or loss events, respectively, and used in the by-event analyses. Since there was more than 1 sample available for 43% of the patients studied, we established a by-patient call for each BAC for that patient as follows: a) if all the samples with data for the BAC had the same call, that consensus call was used, b) if any sample had a combination of copy number gain or loss and no change, the call was gain or loss, respectively, c) in the rare (<0.01% of the BACs examined) cases where one sample had a gain and another had a loss, the majority call was used (e.g., 2 samples with loss and one with gain would

be called a loss at that BAC), and d) if all samples from a patient were no data, the BAC was classified as no data.

Data from individual BACs were not used in further analyses if >40% of the BACs in a group of patients (e.g., <HGD) had a call of no data, suggesting poor hybridization for that particular BAC on the array. Data from chromosomes X and Y were not included in further analyses since a common male DNA was used as a normal control, making gains and losses on these chromosome difficult to quantify for all samples. Any BACs that showed a pattern of alterations that correlated significantly with a particular manufacturing batch (t-test with $p<0.05$ between different manufacturing batches) were considered artifacts and were not included in the analyses (29 total).

Identification of significant copy loss and gains and comparison among different progression stages. For the largest sub-group of patients in this study (72 <HGD patients) to have 99% confidence that loss at a given BAC is significantly different than no loss (null hypothesis), the cutoff is 7 patients, which corresponds roughly to 10% of the patients examined (Fisher's exact test). Therefore, cutoffs on the figures and in our analysis were set at 10%. Due to the higher frequency of alterations in the EA samples, an arbitrary cutoff of 40% was used to identify those alterations that were most frequent in the EA samples. Tukey's test was used to evaluate associations between the mean numbers of alterations present at different stages of progression. The amount of the genome affected by CNAs was calculated by summing the size of regions affected by gains and losses for each non-EA patient; if a patient had more than one sample, the

sample with the greatest amount of the genome affected was used for subsequent analyses. Cox regression model was used to determine if there was a significant relationship between total CNA size and the development of either a DNA content abnormality or EA at a later time point during patient follow-up. As well, Cox regression analysis was used to identify BACs with CNAs associated with development of DNA content abnormalities or EA during follow-up.

RESULTS

Characteristics of the cohort are shown in Table 1. We first examined the frequency of copy number alterations in patients without HGD, in those with HGD, and in those with EA. Examples of representative CNAs are shown in Supplemental Figure 1. Evidence of chromosomal instability, as assessed by the percent of BACs with a copy number alteration, increased significantly in samples from patients without HGD (1.3%), to those with HGD (4.7%), to EA (30.4%) (Supplemental Table 1) ($p<0.0001$, Tukey's test).

We observed different chromosome instability patterns in the frequency and size of CNAs across the spectrum of progression in BE. Throughout the genome, patients with more advanced histology (HGD, EA) had a greater number of copy number change events (any contiguous region of the genome having the same copy number change) and the events were larger than in <HGD. There was a significant increase in number of CNA loss events as well as increased size of those events when comparing patients with <HGD, HGD and EA (Table 2a) ($p<0.0001$ for all comparisons). We found similar results when we examined loss events at a specific locus, p16/CDKN2a/ARF on

chromosome 9p (Table 2b and Supplemental Figure 2). The same CNAs were observed in multiple samples from the same patient across as much as 8cm of the BE segment in the esophagus (data not shown), indicating clones with CNAs undergo clonal expansion similar to other types of alterations[40].

DNA content flow cytometric abnormalities are manifestations of chromosomal instability in many types of cancer and they have been reported to carry an increased risk for progression from BE to EA[26, 41, 42]. We examined the relationship between the number of BAC alterations and DNA content ploidy for each of the 98 patients in this study (Figure 2). The median number of BAC alterations in patients with a DNA content aneuploid population was significantly higher than those with only diploid cell populations (1275 vs 24.5, $p<0.0001$). The vast majority of the diploid samples (141/155; 91%) had less than 180 BAC alterations, compared to 0/19 aneuploid samples. Using an empirical thresholding method, we determined a threshold of 760 BACs with CNAs would allow identification of aneuploid samples with a sensitivity and specificity of 93% and 98%, respectively. Results from bootstrap analysis showed a robust threshold range, with thresholds from 200 to 800 BAC alterations for the identification of aneuploid samples leading to mean sensitivities of 84.0% to 94.8% and specificities of 92.2% to 99.4%, respectively. We quantified the relationship between total number of BAC alterations and probability of being aneuploid (p) with logistic regression $p = \frac{1}{1+e^{-(-19.34+2.09*\log_2(N+1))}}$, where N is total number of BAC alterations per sample (95% CI for the two parameters, -18.6 to -20.1 and 2.0 to 2.2, respectively). This model predicts aneuploidy accurately using the overall number of BACs displaying CNAs.

We then examined genome-wide assessment of copy number abnormalities as a measure of chromosome instability for patient risk assessment for progression to EA or validated intermediate endpoints. Patients whose biopsies contained copy number alterations involving more than 70 Mbp of the genome had a significantly increased risk of progressing to DNA content abnormalities or EA during follow-up (HR=4.9, 95% CI 1.6-14.8, p=0.0047), and the risk increased as more of the genome was affected.

The most common region of copy number alteration in patients without HGD or EA was loss in and around the p16 locus on chromosome 9p (42.9%), along with two other areas distinct from p16 on chromosome 9p: from 10.4 Mb to 11.8 Mb (18.3%) and from 25.5 Mb to 27.5 Mb (19.4%) (Figure 1a and Table 3a). Losses were also observed around 185Mb on chromosome 1q (37.5% of patients), and at 101 Mb on chromosome 8 (41.2%). Other losses at single BACs at frequencies of 10% or more in <HGD patients are listed in Table 3a. The most frequent gains in the <HGD patients involved the very ends of the p arms of chromosomes 17 and 18 (29.2% of patients for each), and gains involving predominantly whole chromosomes were observed on chromosomes 8 (in four patients) and 18 (six patients).

Copy number alterations were more common in patients with HGD and involved larger regions of the genome (Figure 1b and Table 3a). The region in and around the p16 locus was again lost in a large fraction of the patients (45.5%), but losses were observed in more than 10% of patients involving large regions of chromosomes 1, 2, 5, 6, 7, 8, 9, 10,

11, 13, 14, 16, 17, 18, and 22. Gains were seen on chromosomes 3, 5, 8, 14, 16, 17, 18, and 19. Some regions, such as chromosome 5p, 8q, 14q, and 17p showed amplification in one subset of patients and deletion in another (Table 3a). We used Cox regression analysis to identify regions of the genome significantly associated with future development of DNA content abnormalities or EA (Table 4). While the number of EA and DNA content abnormality events in this cohort were small (8 EA and 16 DNA content events out of 71 patients with follow-up data), these data indicate genomic regions that may be of interest in future biomarker studies.

The copy number alterations observed in EA patients indicate accumulation of complex, multiple amplification and deletion events (Figure 1c and Table 3a). All samples from these patients were aneuploid by flow cytometry. The high frequency and large average size of alterations in the EA samples makes it difficult to identify individual gene alterations that may be required for progression to cancer; however, we have listed the regions with most frequent copy number alterations (occurring in at least 40% of the patients) along with potential genes of interest in those regions in the EA patients in Table 3a. High-level amplification events were observed only in EA patients and in a single HGD patient who subsequently progressed to EA (Table 3b).

DISCUSSION

Our study advances validation of chromosome instability as a biomarker for risk assessment in BE by demonstrating for the first time that array CGH can be used as a

common platform to assess chromosomal instability as a predictor of progression in BE. The current standard for risk stratification for patients with BE, dysplasia classification, has several limitations, including observer variation in diagnosis and requirements for large numbers of biopsies[29, 30, 34]. In fact, even what constitutes the histologic definition of Barrett's esophagus is a matter of ongoing debate[43, 44]. In this prospective study, we have examined samples from a cohort of patients representing the spectrum of BE, including both high-risk patients that progressed to EA at a later time point and low-risk patients who did not develop EA, in some cases for almost 12 years of follow up. We have shown that array CGH provides a common platform for assessing genome-wide and locus specific chromosomal instability compared to previous platforms that required combined STR analysis of LOH and DNA content by flow cytometry[26]. We have demonstrated in this cohort study that array CGH can assess genome-wide chromosome instability, like the previously validated biomarker DNA content flow cytometry, and that array CGH can be used to detect patients at increased risk for progression to validated intermediate endpoints such as DNA content abnormalities and EA.

Somatic CNAs are thought to occur rarely in non-neoplastic tissues, and the high frequency of their occurrence across the spectrum of cancer types indicates that loss of genome integrity plays an important role in neoplastic progression. The use of a genome wide measure of genetic instability (CNAs in this study) is appealing since all cancers progress through some type of genetic instability (reviewed recently in[45, 46]). While some cancers may display little overall copy number instability, e.g., MIN cancers, these generally represent a minority of solid tumors, and certainly a minority of EAs[47]. Flow

cytometric analysis of ploidy has been a validated standard for determining gross chromosomal instability, and aneuploid or tetraploid populations are associated with increased risk of EA in patients with BE[26, 28, 42], yet differences in DNA content greater than 10% compared to normal cells (equivalent to ~300Mbp) are required before a flow cytometric determination of aneuploidy can be made confidently. Our results indicate that array CGH is able to identify patients with a significantly increased risk of progression when only 70Mbp of the genome was involved in CNAs, which is less than one-quarter of the changes required by flow cytometry. These results were obtained using only a few samples from each patient - in some cases only a single biopsy from an 11-cm Barrett's segment. Since we know multiple clones can exist in a BE segment, one biopsy every two cm sampling of the segment as reported by Galipeau et al[26] is likely to improve the determination of patient risk. As well, the use of SNP arrays that can measure both LOH and CNAs at a much higher density than BAC arrays would be the most direct means of extending this study to a larger number of patients and testing its utility in the clinic.

The data obtained from this cohort study allow us to identify and examine potentially interesting regions of the genome undergoing CNAs in patients at different stages of progression, extending the findings from earlier pilot studies that examined patients with primarily advanced disease, and did not evaluate the utility of a measure of chromosomal instability as an indicator of progression risk[24, 33]. We found 9p loss encompassing p16 throughout progression, losses on chromosome 5q, 13q and 18q in HGD and EA and high level amplification at ErbB2 on chromosome 17q in EA patients, all of which have been previously identified using different approaches by multiple investigators[21, 48]. While

localized loss of p16 may be too frequent in early BE to be a discriminator of progression risk, these other alterations, as well as expansion of 9p losses to regions beyond the p16 locus, may be robust components of a chromosome instability array platform for further validation in future biomarker validation studies (see also Table 4). Two regions of the genome that have been frequently reported as altered in BE are the FHIT locus on chromosome 3p and the TP53 locus on chromosome 17p[24, 33]. We did not detect FHIT alterations since there was no BAC spanning the locus on our arrays, and the frequency of loss events at TP53 was just below the threshold for reporting (10% of HGD, 33% of EA patients). However, loss of heterozygosity of TP53 can involve copy neutral mechanisms and/or copy gain in nearly 70% of cases[49], so simple copy loss assessment likely under represents the frequency of chromosome instability at this locus.

We also detected examples of clones with mutually exclusive CNA events (i.e., amplification in one patient, loss in another) that can be selected at different points during progression. One example is the prostaglandin-endoperoxide synthase 2 (PTGS2 or COX2) gene, which is overexpressed in a wide variety of cancers[50-52]. We found amplification of COX2 in 27% of the EA cases, but also observed copy loss of COX2 in 37% of <HGD. It is possible that the environment of the reflux exposed esophagus, with its associated chronic inflammation, selects for loss of the COX2 gene. There was a trend for fewer patients with deletion in the region of COX2 to develop DNA content flow cytometric abnormalities during follow-up (4/29, 14%) compared to those lacking the deletion (12/43, 28%), although the difference in this study was not significant. A recent meta-analysis of COX2 expression in BE and EA[53] concluded that there was conflicting

evidence over the role of COX2 in neoplastic progression in BE; our finding of a subset of patients having a deletion in the COX2 locus may explain this heterogeneity in previous studies.

Previous studies that examined primarily EA samples[9-21] reported widespread CNAs throughout the genome, and those that examined a small number of BE samples[22, 23] found far fewer alterations at earlier stages. The previous study by Lai et al[24], using high density Affymetrix arrays, demonstrated alterations within a patient can become more frequent and larger during disease progression, but only examined six highly selected patients that had developed specific genetic alterations. The most recent study by Li, et al, a pilot discovery study using a 33K SNP array to investigate LOH and CNAs in 34 primarily high-risk patients with BE and 8 patients with EA, also found increased CNAs in later stages of progression and an association between number of alterations and aneuploidy[33]. The study presented here extends these earlier observations by demonstrating that a genome wide measure of CNAs can be used as a measure of risk of progression to DNA content abnormalities or EA in a prospective cohort study. Genome-wide arrays have potential for providing accurate cancer risk assessment using a single platform in patients with BE and represents an advanced stage of validation for chromosome instability as a biomarker of cancer risk ready for further validation in larger patient cohorts with prolonged follow-up[25].

The translation of biomarkers identified in discovery studies to a clinical setting requires demonstrating the utility of a biomarker for assessing risk of progression in prospective

cohort studies and adapting the biomarkers to platforms that can be standardized for clinical use. The biomarker panel of 9p LOH, 17p LOH and DNA content that was validated in a 10 year prospective study is able to identify patients at both high and low risk for developing EA, but requires short tandem repeat polymorphisms for assessing LOH and DNA content flow cytometry to detect ploidy alterations, both of which were state of the art when the study was designed in the mid 1990s[26]. As we report here, advancing array technology now can provide a common platform for detecting chromosome instability that is able to detect aneuploid populations, identify patients at risk for future development of ploidy alterations or EA, and identify specific chromosomal regions that undergo frequent CNAs as candidates for additional evaluation.

Literature Cited

1. Cook MB, Wild CP, Everett SM, et al. Risk of mortality and cancer incidence in Barrett's esophagus. Cancer Epidemiol Biomarkers Prev 2007;16:2090-6.
2. Pohl H, Welch HG. The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. J Natl Cancer Inst 2005;97:142-6.
3. Brown LM, Devesa SS. Epidemiologic trends in esophageal and gastric cancer in the United States. Surg Oncol Clin N Am 2002;11:235-56.
4. Thomas T, Abrams KR, De Caestecker JS, Robinson RJ. Meta analysis: Cancer risk in Barrett's oesophagus. Aliment Pharmacol Ther 2007;26:1465-77.
5. Anderson LA, Murray LJ, Murphy SJ, et al. Mortality in Barrett's oesophagus: results from a population based study. Gut 2003;52:1081-4.
6. Gebhart E. Comparative genomic hybridization (CGH): ten years of substantial progress in human solid tumor molecular cytogenetics. Cytogenet Genome Res 2004;104:352-8.
7. Balsara BR, Testa JR. Chromosomal imbalances in human lung cancer. Oncogene 2002;21:6877-83.
8. Schlomm T, Kirstein P, Iwers L, et al. Clinical significance of epidermal growth factor receptor protein overexpression and gene copy number gains in prostate cancer. Clin Cancer Res 2007;13:6579-84.

9. Albrecht B, Hausmann M, Zitzelsberger H, et al. Array-based comparative genomic hybridization for the detection of DNA sequence copy number changes in Barrett's adenocarcinoma. J Pathol 2004;203:780-8.
10. Doak SH, Jenkins GJ, Parry EM, et al. Chromosome 4 hyperploidy represents an early genetic aberration in premalignant Barrett's oesophagus. Gut 2003;52:623-8.
11. El-Rifai W, Frierson HF, Jr., Moskaluk CA, et al. Genetic differences between adenocarcinomas arising in Barrett's esophagus and gastric mucosa. Gastroenterology 2001;121:592-8.
12. Menke-Pluymers MB. Risk factors for neoplastic progression in Barrett's mucosa. Eur J Surg Oncol 1996;22:311-6.
13. Miller CT, Moy JR, Lin L, et al. Gene amplification in esophageal adenocarcinomas and Barrett's with high-grade dysplasia. Clin Cancer Res 2003;9:4819-25.
14. Moskaluk CA, Hu J, Perlman EJ. Comparative genomic hybridization of esophageal and gastroesophageal adenocarcinomas shows consensus areas of DNA gain and loss. Genes Chromosomes Cancer 1998;22:305-11.
15. van Dekken H, Geelen E, Dinjens WN, et al. Comparative genomic hybridization of cancer of the gastroesophageal junction: deletion of 14Q31-32.1 discriminates between esophageal (Barrett's) and gastric cardia adenocarcinomas. Cancer Res 1999;59:748-52.
16. van Dekken H, Vissers K, Tilanus HW, et al. Genomic array and expression analysis of frequent high-level amplifications in adenocarcinomas of the gastro-esophageal junction. Cancer Genet Cytogenet 2006;166:157-62.
17. van Dekken H, Wink JC, Vissers KJ, et al. Genomic analysis of early adenocarcinoma of the esophagus or gastroesophageal junction: tumor progression is associated with alteration of 1q and 8p sequences. Genes Chromosomes Cancer 2006;45:516-25.
18. Varis A, Puolakkainen P, Savolainen H, et al. DNA copy number profiling in esophageal Barrett adenocarcinoma: comparison with gastric adenocarcinoma and esophageal squamous cell carcinoma. Cancer Genet Cytogenet 2001;127:53-8.
19. Weiss MM, Kuipers EJ, Hermsen MA, et al. Barrett's adenocarcinomas resemble adenocarcinomas of the gastric cardia in terms of chromosomal copy number changes, but relate to squamous cell carcinomas of the distal oesophagus with respect to the presence of high-level amplifications. J Pathol 2003;199:157-65.
20. Walch AK, Zitzelsberger HF, Bruch J, et al. Chromosomal imbalances in Barrett's adenocarcinoma and the metaplasia-dysplasia-carcinoma sequence. Am J Pathol 2000;156:555-66.
21. Nancarrow DJ, Handoko HY, Smithers BM, et al. Genome-wide copy number analysis in esophageal adenocarcinoma using high-density single-nucleotide polymorphism arrays. Cancer Res 2008;68:4163-72.
22. Croft J, Parry EM, Jenkins GJ, et al. Analysis of the premalignant stages of Barrett's oesophagus through to adenocarcinoma by comparative genomic hybridization. Eur J Gastroenterol Hepatol 2002;14:1179-86.
23. Riegman PH, Vissers KJ, Alers JC, et al. Genomic alterations in malignant transformation of Barrett's esophagus. Cancer Res 2001;61:3164-70.

24. Lai LA, Paulson TG, Li X, et al. Increasing genomic instability during premalignant neoplastic progression revealed through high resolution array-CGH. Genes Chromosomes Cancer 2007;46:532-42.
25. Sullivan Pepe M, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst 2001;93:1054-61.
26. Galipeau PC, Li X, Blount PL, et al. NSAIDs modulate CDKN2A, TP53, and DNA content risk for progression to esophageal adenocarcinoma. PLoS Med 2007;4:e67.
27. Albertson DG, Pinkel D. Genomic microarrays in human genetic disease and cancer. Hum Mol Genet 2003;12 Spec No 2:R145-52.
28. Rabinovitch PS, Longton G, Blount PL, Levine DS, Reid BJ. Predictors of progression in Barrett's esophagus III: baseline flow cytometric variables. Am J Gastroenterol 2001;96:3071-83.
29. Montgomery E, Bronner MP, Goldblum JR, et al. Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. Hum Pathol 2001;32:368-78.
30. Reid BJ, Haggitt RC, Rubin CE, et al. Observer variation in the diagnosis of dysplasia in Barrett's esophagus. Human Pathology 1988;19:166-78.
31. Reid BJ, Levine DS, Longton G, Blount PL, Rabinovitch PS. Predictors of progression to cancer in Barrett's esophagus: baseline histology and flow cytometry identify low- and high-risk patient subsets. Am J Gastroenterol 2000;95:1669-76.
32. Schnell TG, Sontag SJ, Chejfec G, et al. Long-term nonsurgical management of Barrett's esophagus with high-grade dysplasia. Gastroenterology 2001;120:1607-19.
33. Li X, P.C. G, C.A. S, et al. Single Nucleotide Polymorphism–Based Genome-Wide Chromosome Copy Change, Loss of Heterozygosity, and Aneuploidy in Barrett's Esophagus Neoplastic Progression. Can Prev Res 2008;1:413-23.
34. Wang KK, Sampliner RE. Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. Am J Gastroenterol 2008;103:788-97.
35. Loo LW, Grove DI, Williams EM, et al. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. Cancer Res 2004;64:8541-9.
36. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature 2006;444:444-54.
37. Iafrate AJ, Feuk L, Rivera MN, et al. Detection of large-scale variation in the human genome. Nat Genet 2004;36:949-51.
38. Yang YH, Dudoit S, Luu P, Speed TP. Normalization for cDNA microarray. In: M. Bittner, Y. Chen, A. N. Dorsel, and E. R. Dougherty, editor. Microarrays: optical technologies and informatics. Bellingham, WA: SPIE, Society for Optical Engineers; 2001. p. 141-52.
39. Hsu L, Self SG, Grove D, et al. Denoising array-based comparative genomic hybridization data using wavelets. Biostatistics 2005;6:211-26.
40. Maley CC, Galipeau PC, Li X, Sanchez CA, Paulson TG, Reid BJ. Selectively Advantageous Mutations and Hitchhikers in Neoplasms: p16 Lesions Are Selected in Barrett's Esophagus. Cancer Res 2004;64:3414-27.

41. Reid B, Blount P, Rabinovitch P. Biomarkers in Barrett's Esophagus: A Guideline for Clinicians. Gastrointest Endosc Clin N Am 2003;13:369-97.
42. Teodori L, Gohde W, Persiani M, et al. DNA/protein flow cytometry as a predictive marker of malignancy in dysplasia-free Barrett's esophagus: thirteen-year follow-up study on a cohort of patients. Cytometry 1998;34:257-63.
43. Vakil N, van Zanten SV, Kahrilas P, Dent J, Jones R. The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. Am J Gastroenterol 2006;101:1900-20; quiz 43.
44. Sharma P, McQuaid K, Dent J, et al. A critical review of the diagnosis and management of Barrett's esophagus: the AGA Chicago Workshop. Gastroenterology 2004;127:310-30.
45. Geigl JB, Obenauf AC, Schwarzbraun T, Speicher MR. Defining 'chromosomal instability'. Trends Genet 2008;24:64-9.
46. Ricke RM, van Ree JH, van Deursen JM. Whole chromosome instability and cancer: a complex relationship. Trends Genet 2008;24:457-66.
47. Falkenback D, Johansson J, Halvarsson B, Nilbert M. Defective mismatch-repair as a minor tumorigenic pathway in Barrett esophagus-associated adenocarcinoma. Cancer Genet Cytogenet 2005;157:82-6.
48. Jenkins GJ, Doak SH, Parry JM, D'Souza FR, Griffiths AP, Baxter JN. Genetic pathways involved in the progression of Barrett's metaplasia to adenocarcinoma. Br J Surg 2002;89:824-37.
49. Wongsurawat VJ, Finley JC, Galipeau PC, et al. Genetic mechanisms of TP53 loss of heterozygosity in Barrett's esophagus: implications for biomarker validation. Cancer Epidemiol Biomarkers Prev 2006;15:509-16.
50. Mann JR, DuBois RN. Cyclooxygenase-2 and gastrointestinal cancer. Cancer J 2004;10:145-52.
51. Zha S, Yegnasubramanian V, Nelson WG, Isaacs WB, De Marzo AM. Cyclooxygenases in cancer: progress and perspective. Cancer Lett 2004;215:1-20.
52. Mobius C, Stein HJ, Spiess C, et al. COX2 expression, angiogenesis, proliferation and survival in Barrett's cancer. Eur J Surg Oncol 2005;31:755-9.
53. Mehta S, Boddy A, Johnson IT, Rhodes M. Systematic review: Cyclo-oxygenase-2 in human oesophageal adenocarcinogenesis. Aliment Pharmacol Ther 2006;24:1321-31.

Acknowledgements: We would like to thank Gao Zhang for creating data sets for wavelets analysis, Dave Cowan for database management, Valerie Cerera, Christine Karlsen, and Kamran Ayub for sample procurement and processing, Patricia Galipeau for manuscript review, the FHCRC Array facility for assistance with BAC array processing, and the participants in the Seattle Barrett’s esophagus cohort whose participation makes this study possible.

Tables and Table Legends

	<HGD	HGD	EA
Number of patients	72	11	15
Mean age	60.3	68.1	63.2
Segment length mean	6	8	NA
Segment length range	<1 to 20	<1 to 19	NA
Male:Female	54:18	10:1	15:0
Number of patients with follow up	61	10	NA
Flow Abnormality at Baseline	3	5	NA
Progression to flow abnormality	13	3	NA
Progression to EA	1	7	NA
Mean follow-up time (months)	90.9	54.8	NA
Range of follow-up times (months)	5.8 to 139.5	15.2 to 131.2	NA

Table 1. Cohort characteristics. Age indicates patient age at time of endoscopy when biopsy examined by BAC was obtained. Genetic alterations, flow abnormalities, and progression to EA and flow abnormalities are listed by patient at the time of baseline endoscopy.

2a.

	Average loss event size (bp)	Mean number of loss events (# per sample)	Average gain event size (bp)	Mean number of gain events
<HGD	465,624	11.6	6,220,314	3.9
HGD	4,150,930	13.7	6,523,304	4.0
EA	13,794,970	30.9	10,658,819	32.2

2b.

	Number of patients with p16 loss	Total samples with p16 loss	Average loss size in bp	
<HGD	32 (44.4%)	60 (45.4%)	1,537,015	p=0.01
HGD	5 (45.5%)	9 (36%)	11,385,776	p=0.07
EA	10 (66.7%)	10 (58.9%)	49,905,670	

Table 2. Average size of loss and gain events. a) Contiguous loss and gain events throughout the genome determined by patient. All differences between categories for losses (<HGD vs. HGD, <HGD vs. EA, HGD vs. EA) were significant ($p<0.0001$). **b)** Contiguous loss and gain events including BACs spanning the p16 coding region determined by patient. Significance values for comparing average loss size between the different categories are indicated. The value for comparison between <HGD and EA is $p<0.0001$.

3a.

	Chromosome	Start	End	Percent patients with Gain	Percent patients with Loss	Genes in region
<HGD	1	163,432,534	163,614,111	1.6	24.2	LMX1A
	1*	184,717,073	186,846,060	1.4	37.5	PTGS2 (COX2), PLA2G4A
	8	74,754,218	74,913,326	1.6	17.5	STAU2
	8	101,279,027	101,431,772	2.0	41.2	SPAG1, RNF19A
	9	10,400,067	11,764,217	0	18.3	PTPRD
	9*	21,210,771	22,953,086	0	42.9	MTAP, CDKN2A, CDKN2B, others
	9*	25,425,786	27,582,822	0	19.4	TUSC1, PLAA, IFT74, TEK
	10	108,165,938	109,275,992	0	12.7	SORCS1
	10	114,856,262	115,024,106	0	28.1	TCF7L2
	11	474,042	626,401	0	12.7	RNH1, HRAS, RASSF7, MUPCDH, SCT and others
	17	2,912,016	3,092,354	2.8	15.3	olfactory receptor genes
	19*	8,714,331	8,864,039	0	12.5	MBD3L1, MUC16
	22*	22,795,270	22,989,794	0	18.0	CABIN1, GGTLA1
	17*	**836,330**	**1,008,128**	**29.2**	**0**	**ABR, TIMM22**
	18	**168,383**	**75,965,502**	**29.2**	**1.5**	**BCL2, MADH2, MADH4, DCC, DPC4, PI5, others**
HGD	1	46,425,202	46,624,292	0	45.5	RAD54L
	1*	184,717,073	186,846,060	0	27.3	PTGS2 (COX2), PLA2G4A
	5	50,107,903	180,611,420	0	27.3	PLK2, CCNB1, XRCC4, APC, RAD50, others
	7*	156,680,451	158,620,885	0	18.2	PTPRN2, NCAPG2
	8	38,227,281	38,404,631	0	18.2	PPAPDC1B, WHSC1L1, FGFR1
	9	222,268	38,427,295	0	45.5	MTAP, CDKN2A, CDKN2B, others
	10	109,092,539	133,471,230	0	18.2	BUB3, C10orf119, NANOS1, others
	11	149,520	6,642,613	0	36.4	MUC6, RHOG, others
	11	8,555,485	8,795,697	0	18.2	STK33, ST5
	11	120,168,346	133,686,875	0	18.2	CHEK1, TIRAP, ETS1, others
	14	58,681,498	106,175,506	9.1	36.4	MLH3, MAX, FOXN3, others
	16	77,215,302	77,345,302	0	18.2	WWOX
	17	18,863,807	19,044,654	0	36.4	GRAP
	18	600,984	771,970	9.1	27.3	TYMS, ENOSF1, YES1
	18	17,908,960	76,089,909	0	36.4	SMAD2, SMAD4, SMAD7, DCC, others
	8*	**7,156,823**	**7,328,299**	**18.2**	**9.1**	**DEFB103A**
	14	**22,278,370**	**23,705,612**	**18.2**	**9.1**	**MMP14, PCK2, others**
	18	**17,274,438**	**18,073,471**	**18.2**	**9.1**	**MIB1, microRNAs**
	19*	**48,859,097**	**50,122,986**	**27.3**	**0**	**ERCC2, multiple zinc finger proteins**
	5	***561,584***	***43,795,937***	***28.6***	***54.5***	***POLS, NDUFS6, SLC6A19, others***
	8	***101,279,027***	***101,431,772***	***11.1***	***11.1***	***SPAG1, RNF19A***
	17	***2,912,016***	***3,092,354***	***18.2***	***27.3***	***Olfactory receptor genes***
EA	5	50,107,903	180,611,420	7	64	PLK2, CCNB1, XRCC4, APC, RAD50, others
	8*	304,159	1,524,873	0	53	FBXO25
	8	19,651,026	26,038,850	13	40	ADAM28, LOXL2, FGF17, others
	9	222,268	30,116,164	7	73	MTAP, CDKN2A, CDKN2B, others
	10	214,399	8,156,391	13	47	GATA3, NET1, others
	15*	19,138,465	20,536,973	0	67	TUBGCP5
	16	77,215,302	77,345,302	0	60	WWOX
	17	2,912,016	3,092,354	7	60	
	17	12,002,245	19,044,654	0	47	COX10, FLCN, others
	18	22,533,011	76,089,909	20	67	SMAD2, SMAD4, SMAD7, DCC, others
	21	14,850,741	46,912,065	7	67	ANA, PCNT, TIAM1, others
	22	15,756,122	49,441,620	13	57	BIK, NF2, CHEK2, others
	7	**835,958**	**107,941,302**	**80**	**13**	**EGFR, CDK6, SMURF1, ABCB1, others**
	8	**114,955,242**	**141,809,117**	**60**	**13**	**ASAP1, MYC, WISP1, PTK2, others**
	15	**83,671,081**	**100,021,943**	**47**	**20**	**FES, PRC1, others**
	17	**30,441,739**	**30,442,082**	**53**	**0**	
	18	**17,274,438**	**20,423,414**	**47**	**27**	**GATA6, RBBP8, others**
	19	**33,315,121**	**63,560,213**	**60**	**27**	**CCNE1, CEACAM5, XRCC1, others**
	20	**9,943**	**62,430,362**	**67**	**20**	**ZNF217, TOP1, DNMT3B, PCNA, other**

3b.

Chromosome	High amp start	High amp end	Patients affected	Log2 ratio
6	43,262,021	43,439,935	2	1.53
6	51,252,527	51,423,711	2	2.53
7	53,704,725	53,864,621	3	1.61
7	55,238,304	55,381,999	4	4.41
8	128,612,329	128,822,827	2	2.11
11	25,048,819	25,049,177	4	4.32
16	10,944,014	11,549,689	2	1.63
17	34,979,166	36,924,021	6	3.87
19	34,301,924	34,664,148	2	2.52

Table 3. List of most frequent chromosomal regions of gain or loss in patients with BE or EA. 3a) Most commonly altered regions in patients with BE. Chromosome regions that contain at least one known CNV thought to be present in greater than 10% of populations analyzed are indicated by an asterisk. Normal type indicates regions of loss, bold type indicates regions of gain, italics indicate regions for which different populations having gain or loss were identified. Genes in region is a subjective list of genes in the altered region that may be of interest for additional studies and is not meant to be exhaustive, particularly in larger regions (e.g., chromosome 18). Start and end of altered region is given in basepairs based upon the locations of the BACs that bound the region being altered. Percent of patients with gains and losses represent the maximum percentage within the region listed. **3b) Regions of high level amplification.** High level amplification was defined as having a Log2 ratio > 1.5. All high level amplification was found in EA samples, with the exception of a single patient with HGD, who subsequently went on to develop EA, having the amplification on chromosome 17. Log2 ratio indicates the maximum value observed across all patients with that amplification.

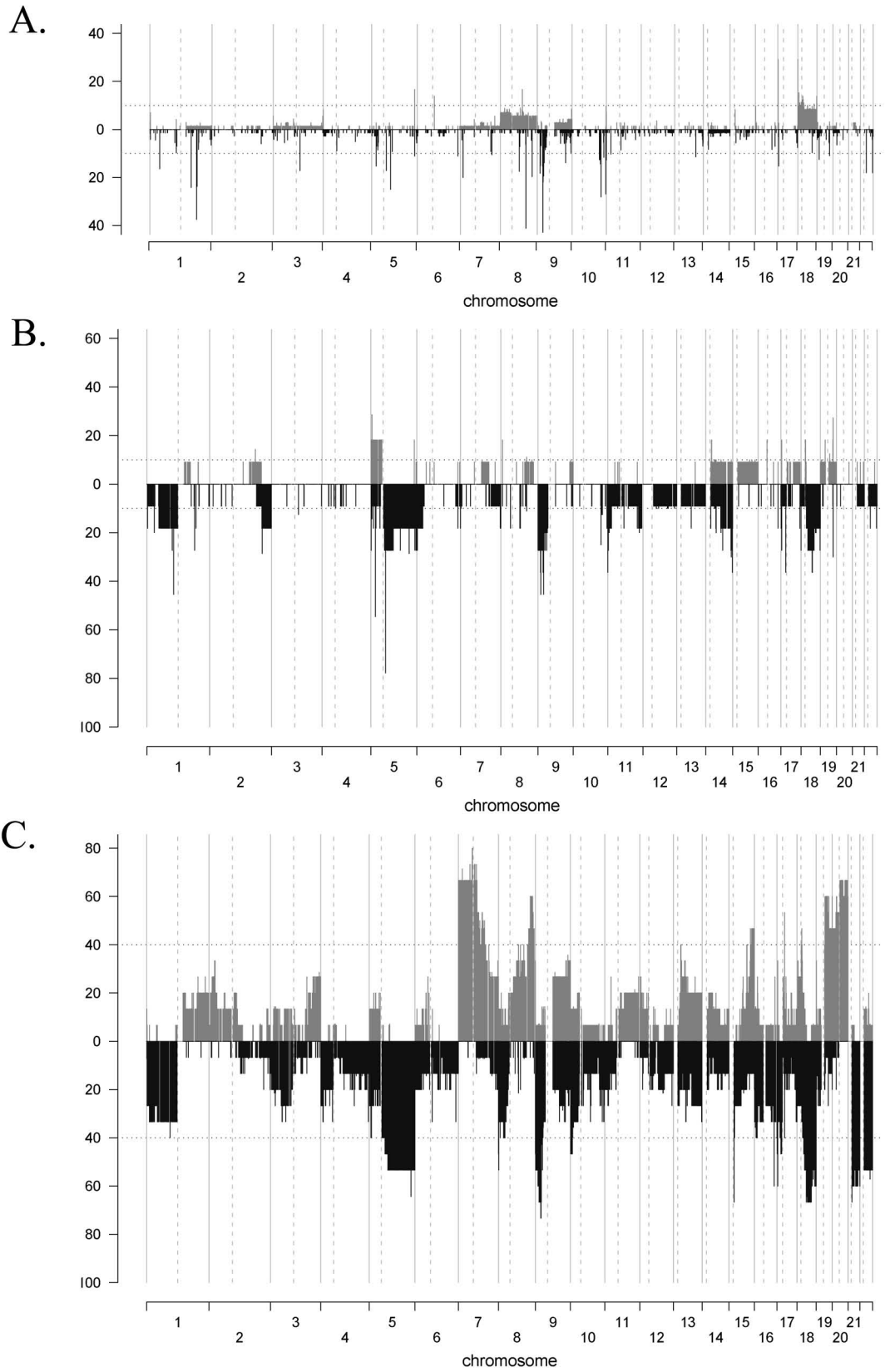
Alteration/ outcome	Chromosome	Region affected (Mbp)	# of BACs	patients with association	p value	RR	Potential genes of interest
Loss/EA	1	98.9 to 104.0	3	4	0.002	10.2	SASS6, CDC14A, COL11A1, others
	5	53.4 to 87.4	42	5	<0.001	20.0	MAP3K1, CCNB1, AGGF1, DMGDH, others
	9	19.7 to 20.0	3	3	<0.001	20.3	SLC24A2
	11	0.5 to 2.9	3	4	<0.001	21.3	CDKN1C, TSSC4, CTSD, LRDD, others
	14	98.9 to 105.2	8	3	<0.001	11.7	BAG5, CKB, MARK3, MEG3
	18	44.0 to 44.8	2	3	<0.001	12.4	SMAD7
Gain/EA	5	0.6 to 43.8	61	3	<0.001	71.0	TPPP, TERT, RAD1, GHR, others
Loss/DNA	6	105.3 to 106.9	3	2	0.005	9.7	HACE1
	9	29.9 to 38.4	16	4	0.003	19.1	NDUFB6, BAG1, SHB
	9	68.8 to 69.2	2	2	0.04	4.9	
Gain/DNA	9	69.8 to 135.8	143	3	<0.001	9.6	CDK9, VAV2, ABL1
	18	0.6 to 24.7	41	7	<0.001	20.2	YES1, TYMS, NDC80, others

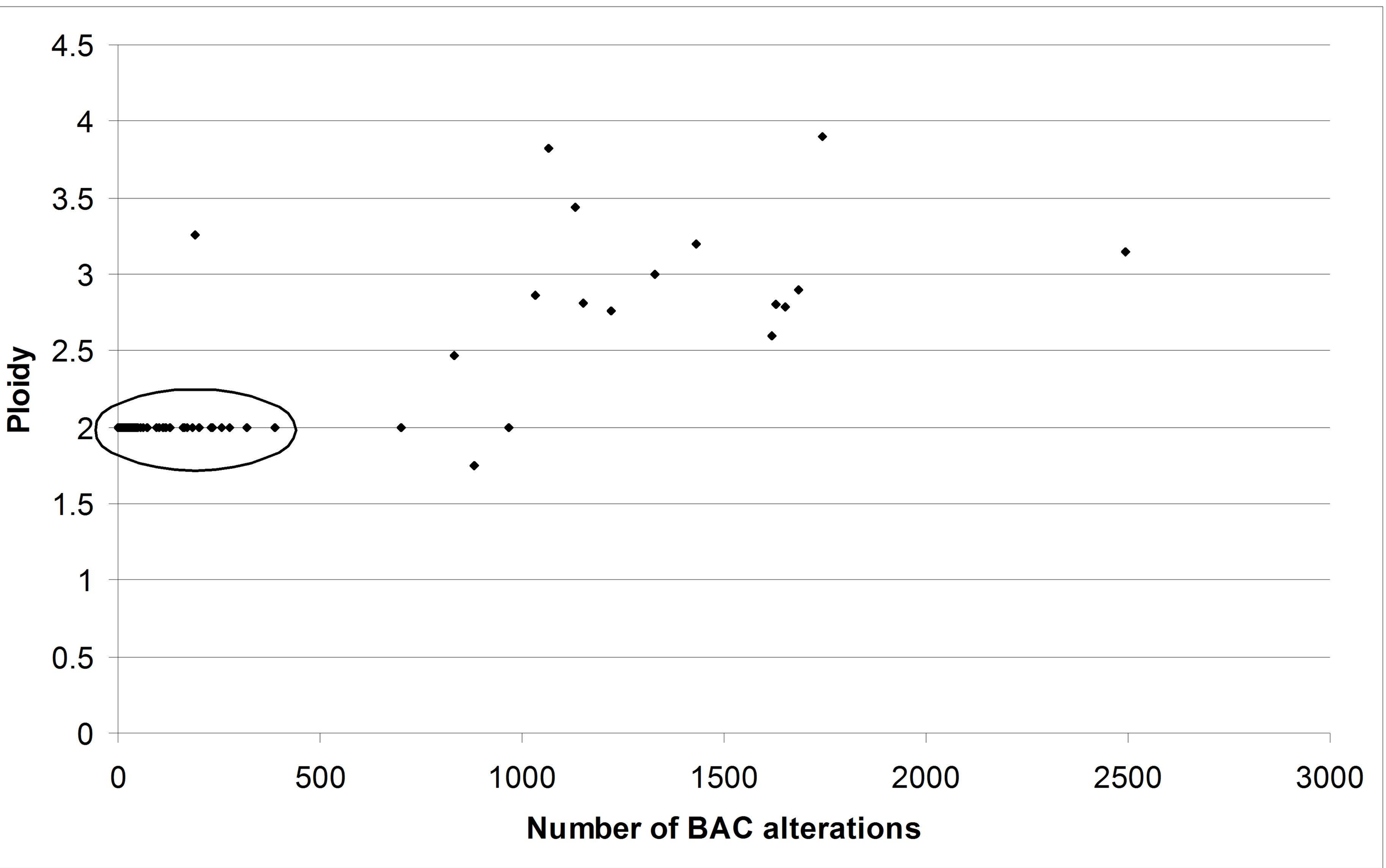
Table 4. Chromosome regions with CNAs associated with future development of EA or DNA content abnormalities. # of BACs indicates how many contiguous BACs are found in the region of interest. Patients with association indicates how many had the alterations indicated out of 16 total follow-up DNA content abnormalities or 8 total follow-up EA cases. Significance and Relative Risk (RR) were determined by Cox regression analysis.

FIGURE LEGENDS

Figure 1a- c. Frequency plots of gains and losses throughout the genome in patients with <HGD (a), HGD (b), and EA (c). Y-axis indicates the percentage of patients having gains (in grey, above 0) or losses (in black, below 0) for each BAC in the array, X-axis indicates position on the chromosome from tip of p-telomere to tip of q-telomere. Dotted vertical lines indicate centromere location. The <HGD and HGD plots have a line at 10% and the EA plot at 40% to identify chromosome regions with most frequent alterations.

Figure 2. Overall number of copy number alterations in diploid and aneuploid samples associate with aneuploidy as measured by flow cytometry. Sample with maximum ploidy, or for diploid patients, with maximum number of BAC alterations for each patient is indicated. Area in the oval encompasses data points from 80 patients.


A.
40
20
0
20
40
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
chromosome
B.
60
40
20
0
20
40
60
80
100
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
chromosome
C.
80
60
40
20
0
20
40
60
80
100
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
chromosome


Ploidy
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
0
500
1000
1500
2000
2500
3000
Number of BAC alterations