An Approach to Analysis of Large-Scale Correlations between Genome Changes and Clinical Endpoints in Ovarian Cancer


Abstract

This report describes analyses of associations of genome copy number aberrations in ovarian cancers with clinical features using genome-wide graphical and analytical procedures. These studies show that tumor grade is a better indicator of the extent of genomic progression than stage, that loss of chromosome 16 occurs preferentially in high-grade tumors, and that gains of 3q26-qter, 8q24-qter, and 20q13-qter occur frequently in low-grade and low-stage tumors and thus may be early events in ovarian cancer development. In addition, loss of chromosome 16q24 and a total number of independent genome copy number aberrations >7 are associated with reduced survival duration. The association of loss of 16q24 (D16S3026) with decreased survival duration was confirmed by quantitative PCR. Regions that frequently are abnormal and associated with altered survival duration are strong candidates for higher resolution analysis and gene discovery and may be useful markers for prediction of clinical outcome.

Introduction

Ovarian cancer is the fifth most common cancer in women in the United States, accounting for approximately one-quarter of the cancers arising from the female genital organs. Unfortunately, the survival rate in ovarian cancer is so low that this disease is responsible for approximately half of all gynecological cancer deaths. Low survival results from the fact that early stage ovarian cancer frequently is not detected because it is mostly asymptomatic and because the cancer quickly disseminates throughout the peritoneal cavity.

Patients with advanced ovarian cancer usually are treated with surgery followed by chemotherapy. Treatment planning is directed by various prognostic factors, most notably stage and histological grade. Although histological grade and clinical stage are loosely correlated, high-grade malignancies (G3, G4) may be encountered in stage I disease, and low-grade malignancies may be advanced in stage. Outcome often is poor for patients with high-grade (G3, G4) stage IA/B tumors, stage IC tumors, and stage IIC tumors and for patients with higher stage disease. However, this stratification scheme is far from perfect because some patients with low-grade and low-stage tumors progress, and 10–20% of patients with advanced stage cancers are cured by a combination of surgery and cisplatin-based therapy. We postulate that these differences in outcome have a genetic basis and that comprehensive genomic analyses will reveal characteristics that can be used to predict response to therapy. Genomic aberrations that are associated with adverse outcome then become targets for gene discovery and ultimately targets for therapeutic development.

Specific genetic aberrations already found in ovarian cancers include amplification and/or overexpression of ERBB2 (1), MYC (2), PFK2CA (3), and AKT2 (4) and mutation or down-regulation of TP53 (5, 6), K-RAS (7), LOT1 (8), DOC2 (9), NOE2 (10), OVCA1 (11), and SPARC (12). However, other genes are likely to be involved. Evidence for their existence comes from identification of numerous regions of recurrent abnormality using classical cytogenetics or CGH and from analyses of allelic imbalance. The regions of recurrent abnormality identified in these studies are presumed to encode genes that contribute to ovarian cancer genesis or progression.

Some of these aberrations have been associated with clinical end points such as survival duration, grade, and stage in an attempt to elucidate the nature of ovarian cancer progression and to identify markers that more accurately stratify patients according to clinical outcome. Dodson et al. (13), for example, suggested that LOH at 6p and 17p was an early event and LOH on 13q and 15q was a late events. Rosen et al. (14) presented evidence for significant amplification of FGF3 (INT2) in advanced stage of ovarian cancer, whereas Tavassoli et al. (15) reported LOH on 5q as an early event because of its high incidence in stage I tumors. In addition, our CGH studies (16) of high- and low-grade tumors suggested that copy number increases at 3q26 and 20q13 might be early events. We also reported a significant association between reduced survival duration and total number of abnormalities detected by CGH (16, 17) and loss of 16q (17).

In this report, we extend these studies using a CGH-based, genome-wide approach for identification of genome aberrations in ovarian cancer that are associated with clinical end points. Statistical and graphical methods that aid in visualization of these genome-wide comparisons are presented and used to identify associations with survival duration, grade, and stage. Significant associations with survival duration are further assessed in an expanded set of tumors, using quantitative PCR.

Materials and Methods

Materials. Fresh tumor samples were obtained from 40 patients at Duke University Medical Center and 20 patient samples from the tissue core of the “Biology of Ovarian Cancer” Program Project. The clinical stage of each patient was determined according to the International Federation of Gynecology and Obstetrics Classification. Stage I and II tumors were grouped as low-stage tumors, and stage III and IV tumors were grouped as high-stage tumors. Each tumor was assigned to a histopathological subtype and a grade. Clear-cell carcinomas and tumors of low malignant potential were excluded. Tumor grade was determined based on a modified Broder’s classification. Grade 1 and 2 tumors were grouped as low-grade, and grade 3 and 4 tumors were grouped as high-grade tumors. The clinical and histopathological characteristics of the 60 tumors analyzed in this study are summarized in Table 1.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 This work was supported by NCI Grant CA64602 and Vysis.

Received 12/9/99; accepted 8/16/00.

5382
Note that the 42 tumors analyzed using CGH are a subset of the 60 tumors analyzed using quantitative PCR.

CGH. CGH was performed generally as described previously (16, 18). Briefly, DNA samples were extracted from fresh-frozen tissue specimen and normal lymphocytes by proteinase K digestion and phenol-chloroform extraction, and labeled by nick translation with fluorescein-12-dUTP and Texas red-5-dUTP (DuPont NEN Products, Boston, MA), respectively. Four hundred ng each of the labeled DNAs plus 30 μg of unlabeled human Cot-1 DNA (Life Technologies, Gaithersburg, MD) were denatured and hybridized to denatured metaphase chromosome preparations at 37°C in a moist chamber for 2 days. After hybridization, slides were washed, air-dried, and counterstained with DAPI in an antidote solution.

Analysis of stained CGH preparations was accomplished using QUIPS Genetics Imaging Software (Vysis, Downers Grove, IL). Three-color DAPI, FITC, and Texas red images were acquired using a Zeiss Axioplan microscope equipped with a cooled CCD camera. Six to eight metaphases were analyzed in each preparation. Chromosomes were identified according to their DAPI banding pattern, and green-to-red fluorescence intensity profiles (±1 SD) were calculated for each chromosome type from measurements of ~10 chromosomes.

Significant gains and losses of relative DNA sequence copy number along each of chromosome type were defined as those for which the mean of green-to-red ratio was >1.25 or <0.75 (16), respectively. The total number of CNAs for each sample was determined by summing the number of contiguous regions of increased/decreased copy number flanked by regions in which the mean green-to-red ratio was 1.0. The gain or loss of both arms of a chromosome was counted as one event.

Quantitative PCR. Copy number was assessed at the three loci listed in Table 2, using quantitative PCR analysis as described by Ginztinger et al., because regions containing these loci frequently were found to be abnormal in copy number in the CGH study and/or associated with reduced survival duration. Briefly, test and reference loci were PCR amplified in the presence of TaqMan probes carrying donor (FAM) and acceptor (TAMRA) fluorescence molecules. The amount of donor fluorescence in each reaction liberated by the exonuclease degradation of the TaqMan probe during PCR amplification was measured as an indication of the amount of amplified material. Copy numbers at these loci were determined relative to a pooled reference comprising six microsatellite loci (D1S2868, D2S385, D4S1605, D5S643, D10S586, and D11S1315) selected from regions that showed few copy number abnormalities in the CGH analyses. Primer sequences for these loci were obtained from the Whitehead Institute Center for Genome Research. The RCN at each test locus was defined as:

\[
\text{RCN} = 2 \times 2^{\frac{\Delta Ct(\text{normal}) - \Delta Ct(\text{tumor})}{\Delta Ct(\text{normal})}}
\]

where \(\Delta Ct(\text{normal}) = Ct(\text{test}) - Ct(\text{reference})\) for normal DNA, and \(\Delta Ct(\text{tumor}) = Ct(\text{test}) - Ct(\text{reference})\) for tumor DNA.

\(\Delta Ct(\text{normal})\) was the mean of measurements on 16 unrelated normal DNAs. Standard deviations of \(\Delta Ct(\text{normal})\) were <0.3 for all test loci. The pooled SD for all loci was used to calculate a 95% tolerance interval as described. Samples were scored as abnormal at a test locus if the RCN measured at that locus fell outside of the tolerance interval of 0.73–1.37. Thus, DNA copy numbers >1.37 were scored as having significantly increased RCN, and those <0.76 were scored as having significantly reduced RCN.

Statistical and Graphical Methods. Statistical and graphical methods were developed to aid in evaluating the association of CGH with outcome, grade, and stage. For purposes of statistical analysis, the entire genome was divided into 245 regions of equal length. Thus, there were 20 regions for chromosome 1 compared with 4 regions for chromosome 22. For each tumor, a time to recurrence indicator was defined to be 1 for recurrent tumors and 0 for tumors that had not recurred during the follow-up time (i.e., observation was censored). CGH information also was recorded as an indicator for each region: +1 for losses; +1 for gains; 0 for neither gain nor loss.

The relationship between genome copy number changes and grade or stage was assessed using a 2 × 2 contingency table for each of the 245 regions. Specifically, the numbers of tumors with gain or loss were determined for each region, and the results were assembled into a 2 × 2 table as shown below:

\[
\chi^2 = \frac{(ad - bc)^2N}{(a+b)(c+d)(a+c)(b+d)}
\]

To aid interpretation, \(\chi^2\) values were plotted against distance along genome. A similar method also was applied to survival data. However, survival analyses were accomplished by fitting Cox proportional hazards models to the CGH abnormality indicators in each region. This model was used to test associations between copy number change, as detected by CGH, and survival time. The Cox model does not require that survival follow a parametric form, but assumes that risk of recurrence or death is increased by a constant factor when the factor is present compared with risk when the factor is not present. It also assumes that risk for several factors is equal to the product of the risks for the individual factors. The Cox model was applied to our data using a function written in S-Plus that included the capability to consider other predictive factors, e.g., stage and grade. Thus, the predictive value of a genome aberation could be determined in the presence of other prognostic factors. The function produced a set of p-values corresponding to the statistical significance of gain or loss at each genomic region. These results were then plotted as signed log p-values, where the sign was negative for regions of loss and positive for regions of gain. p-values were plotted only for regions in which >20% of the tumors had genome copy number abnormalities.

Typically, in a large-scale correlative study, there will be several regions with p-values below the p-values = 0.05 level of significance and a few below p-values = 0.01 because of the large number of tests performed. Statisticians call this the “multiple testing problem.” To understand the impact of multiple testing in the present study, we performed a series of 100 random permutation tests in which the relationship between the genomic regions and the outcome of interest (e.g., stage, grade, or survival) was destroyed by permuting the outcomes while leaving the genetic information unperturbed. Results from the permutation tests were analyzed to assess the likelihood that the associations observed in this study had true significance. This approach may also be useful in assessing the significance of associations between outcome and gene expression or gene dosage established for a large number of loci by array analysis technologies.

Results and Discussion

CGH. CGH analyses were performed for 18 low-stage and 24 high-stage ovarian cancers. The genome CNA data are plotted in Fig. 1, a and b, to show differences between low- and high-grade tumors. This comparison shows clear differences. Specifically, the total number of CNAs was significantly higher in high-grade tumors than in low-grade tumors (Mann-Whitney, P = 0.0005), supporting an association established earlier (16). Copy number increases at 3q26-qter, 8q24-qter, and 20q13-qter were frequent in low- and high-grade tumors, suggesting these as early events in ovarian cancer. This confirms the finding by Iwabuchi et al. (16) that gains of 3q and 20q occur early in cancer development. Fig. 1e shows a large \(\chi^2\) value for
loss of 4q, suggesting that this alteration occurs more frequently in high-grade than low-grade tumors. The \( \chi^2 \) statistic for this change (\( \sim 19 \)) was much larger than the largest statistic (13) found during permutation testing.

The frequencies of recurrent aberrations in low- and high-stage tumors are shown in Fig. 1. For high-stage tumors, copy number decreases occurring in \( >40\% \) of low-stage tumors occurred at 3q26-pter, 8q24-pter, 17q25-pter, and 20q13-pter. No copy number decreases occurred in \( >40\% \) of low-stage tumors. Gain of 20q was the most frequent abnormality in low-stage tumors, occurring in \( >75\% \). High-stage tumors showed these same recurrent abnormalities as well as recurrent gains involving 1q22-q25, 1q41-pter, and 7q32-pter, and recurrent losses involving 4q, 13q21-qter, 16q13-qter, 18q21-qter, and 19q13-q25. The overall aberration frequency was not significantly different between low- and high-stage tumors. Fig. 1f shows that none of the differences between low- and high-stage tumors detected by CGH were significantly different from those detected during permutation testing. This suggests that stage is not a good measure of the extent of genome evolution in ovarian cancers.

The total number of CNAs was significantly associated with survival duration as described in earlier studies (16, 17). In the earlier studies, patients with tumors with \( \leq 5 \) CNAs were found to survive significantly longer than patients with tumors with \( >10 \) CNAs. The present study also shows this difference \( (p < 0.001) \). If the number of CNAs is considered as a continuous variable, the association with reduced survival duration is also significant \( (p = 0.02) \), based on Cox proportional hazards model with no other factor. If grade is added, then \( p = 0.057 \) for CNAs as a continuous variable, and grade is not significant \( (p = 0.92) \). In this analysis, the single cut-point that minimizes \( p \) is number of CNAs \( > 7 \) versus number of CNAs \( \leq 7 \), which has a probability \( p = 0.08 \), taking into account not having a priori knowledge of the best cut-point.

We also analyzed associations between survival duration and genome CNAs for all regions of the genome. Fig. 2 shows the \( p \)-values for the associations of genome CNAs with survival duration plotted as a function of abnormality location along the genome. \( p \)-values associated with increased copy number are plotted as positive values, and those associated with decreased copy number are plotted as negative values. Regions of copy number increase associated with reduced survival duration included 1q, 3q, and 7q. Regions of copy number decrease associated with reduced survival duration included 4p, 16q, 18q, and both arms of the X chromosome. The strongest association with reduced survival duration was with loss of 16q. Moreover, this association held even for late-stage patients stratified according to 16q status. As expected, we did find associations with \( p < 0.05 \) during permutation testing, suggesting that some of these associations with survival are likely to have happened by chance. This motivated our efforts to analyze aberrations using an alternative copy number analysis technology.

**PCR Analysis.** Sixty tumors (including the 42 analyzed using CGH) were assessed for copy number at three loci, using quantitative PCR to measure genome copy number by a technology other than CGH. Two loci on 16q23-q24 (D16S3026 and D16S3033) were chosen because of the strong association between survival duration and reduced copy number in this region \( (p = 0.0013) \) and an earlier (16) CGH study. MYC on chromosome 8q24 was chosen because of the importance of this oncogene in numerous cancers and because the region frequently was increased in copy number. Increased MYC copy number and decreased D16S3033 copy number were not significantly associated with survival duration, but decreased copy number at 16q24.2 (D16S3026) was associated with poor sur-
vival. The association of loss at D16S3026 was significant in all tumors \((p = 0.0019)\) and in the 42 tumors tested by CGH \((p = 0.0012)\), confirming the association established using CGH. However, the association was not as significant in high-stage tumors \((p = 0.0799)\) as was found using CGH \((p = 0.0013)\). This same region was significantly associated with reduced survival in an earlier study by Iwabuchi et al. \((16)\). Thus, loss of function of genes in this region is likely to contribute to ovarian cancer progression and may serve as a useful marker of poor outcome. \textit{E-cadherin (CDH1)} is a candidate tumor suppressor gene at 16q22 \((19)\). However, the region most frequently lost in this study was distal to D16S3033 at 16q23 loss of CDH1, and was not associated with reduced survival duration. This suggests that genes other \textit{CDH1} are important in ovarian cancer.

In summary, we have developed a large-scale strategy for association of genotype with clinical behavior and used it to identify several recurrent abnormalities that are associated with tumor grade and clinical outcome. Regions that are frequently abnormal and associated with altered survival duration are strong candidates for higher resolution analysis and gene discovery and may be useful markers for prediction of clinical outcome.

References

An Approach to Analysis of Large-Scale Correlations between Genome Changes and Clinical Endpoints in Ovarian Cancer


*Cancer Res* 2000;60:5382-5385.

Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/60/19/5382

Cited articles
This article cites 18 articles, 10 of which you can access for free at:
http://cancerres.aacrjournals.org/content/60/19/5382.full.html#ref-list-1

Citing articles
This article has been cited by 19 HighWire-hosted articles. Access the articles at:
/content/60/19/5382.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.