High-Resolution Single Nucleotide Polymorphism Array Analysis of Epithelial Ovarian Cancer Reveals Numerous Microdeletions and Amplifications

Kylie L. Gorringe,[1] Sharoni Jacobs,[2] Ella R. Thompson,[1,3] Anita Sridhar,[1] Wen Qiu,[1,3] David Y.H. Choong,[1] and Ian G. Campbell[1,3]

Abstract

Purpose: Genetic changes in sporadic ovarian cancer are relatively poorly characterized compared with other tumor types. We have evaluated the use of high-resolution whole genome arrays for the genetic profiling of epithelial ovarian cancer.

Experimental Design: We have evaluated 31 primary ovarian cancers and matched normal DNA for loss of heterozygosity and copy number alterations using 500K single nucleotide polymorphism arrays.

Results: In addition to identifying the expected large-scale genomic copy number changes, >380 small regions of copy number gain or loss (<500 kb) were identified among the 31 tumors, including 33 regions of high-level gain (>5 copies) and 27 homozygous deletions. The existence of such a high frequency of small regions exhibiting copy number alterations had not been previously suspected because earlier genomic array platforms lacked comparable resolution. Interestingly, many of these regions harbor known cancer genes. For example, one tumor harbored a 350-kb high-level amplification centered on *FGFR1* and three tumors showed regions of homozygous loss 109 to 216 kb in size involving the *RB1* tumor suppressor gene only.

Conclusions: These data suggest that novel cancer genes may be located within the other identified small regions of copy number alteration. Analysis of the number of copy number breakpoints and the distribution of the small regions of copy number change indicate high levels of structural chromosomal genetic instability in ovarian cancer.

Ovarian cancers commonly undergo complex cytogenetic alterations and are often aneuploid (1). Metaphase comparative genomic hybridization (CGH) and allelic imbalance analyses carried out over many years by numerous investigators have enabled the identification of recurrent regions of copy number gain, loss, and/or allelic imbalance. For example, copy number gains are frequently observed on 1q, 3q, 8q, 12p, and 20, as are losses and allelic imbalance on chromosome 4, 13q, 17, and 22q (2–5). However, despite such knowledge, the identification of the specific genes that are the target of these alterations has proven difficult. Although this lack of success could be attributed to the random nature of genomic alterations, an equally valid explanation is that past technologies have lacked the resolution to readily identify the underlying cancer-associated genes.

Recent molecular genetic studies of a variety of tumor types have identified tumors with highly complex chromosome rearrangements and have found that relatively small regions of allelic imbalance and copy number gain may be quite common (6). For example, an array CGH study of breast cancer detected an average of five high-level amplifications under 2 Mb in size per cell line (7). Additionally, a comparison of Affymetrix Mapping 10K array data with 399 microsatellite markers analyzed in lung cancer found that twice as many regions of allelic imbalance were identified using the single nucleotide polymorphism (SNP) array, including multiple small regions not detected by microsatellites; the minimum allelic imbalance region detected dropped from 17.5 Mb in the microsatellite analysis to 1.6 Mb by the 10K SNP array (8). The presence of decreasingly sized aberrations being detected implies a continuum of genetic instability phenotypes from simple copy number changes affecting whole chromosomes through to multiply fractured and rearranged chromosomes.

The highest-resolution ovarian studies to date used bacterial artificial chromosome array CGH with an average 1 Mb spacing to analyze 23 ovarian cancer cell lines (9) and 17 primary tumors (10). However, the majority of genetic studies of ovarian cancer have used low-resolution techniques, such as

Authors' Affiliations: [1]Victorian Breast Cancer Research Consortium Cancer Genetics Laboratory, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia; [2]Affymetrix, Inc., Santa Clara, California; and [3]Department of Pathology, University of Melbourne, Melbourne, Victoria, Australia

Received 2/28/07; revised 4/30/07; accepted 5/16/07.

Grant support: Victorian Breast Cancer Research Consortium, Australia. E.R. Thompson is a recipient of a National Breast Cancer Foundation Postgraduate Scholarship. W. Qiu is a recipient of an Australian National Health and Medical Research Council Dora Lush Postgraduate Scholarship.

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.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Requests for reprints: Ian G. Campbell, Victorian Breast Cancer Research Consortium Cancer Genetics Laboratory, Peter MacCallum Cancer Centre, Locked Bag 1, A'Beckett Street, Melbourne, Victoria 8006, Australia. Phone: 61-3-96561803; Fax: 61-3-96561411; E-mail: ian.campbell@petermac.org.

doi:10.1158/1078-0432.CCR-07-0502

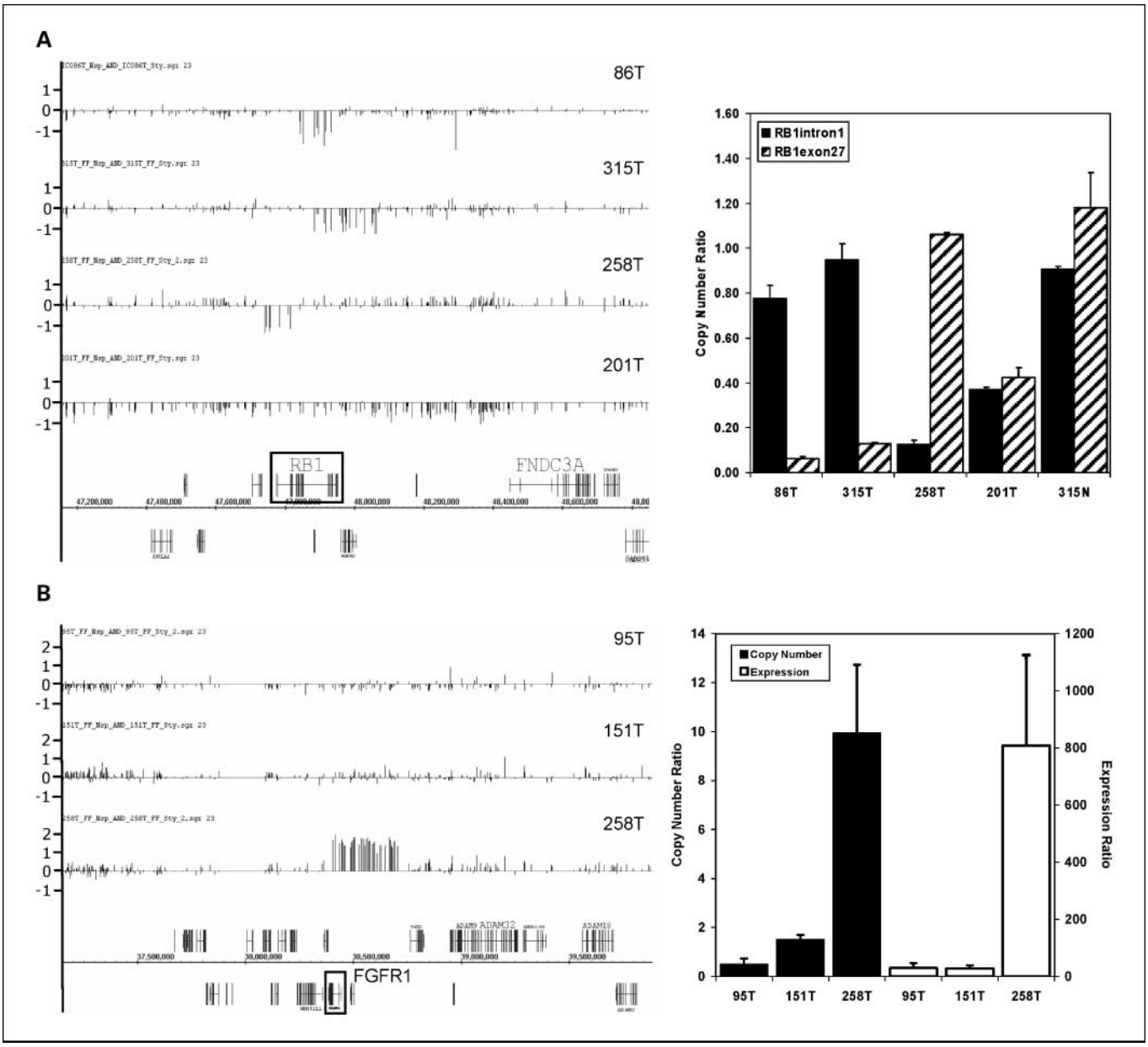

Fig. 1. Small regions of copy number change over known cancer genes. *A, RB1*. Left, SNP $\log_2$ values of three tumors with homozygous deletions over *RB1* on chromosome 13. Tumor sample 201T, also shown, has a large hemizygous deletion spanning most of chromosome 13. Each bar represents the $\log_2$ intensity ratio of a single SNP relative to the matched normal sample, with bars above and below the midline showing increased and decreased hybridization intensity, respectively. Confirmation of the tumor deletions by QPCR of *RB1* using two sets of primers, one at either end of the gene, is shown at right. Also shown is a normal sample (315N). *B, FGFR1*. Left, SNP $\log_2$ values of three tumors, one with a 350-kb amplification over *FGFR1*. Confirmation of the *FGFR1* amplification and increase in expression in sample 258T by QPCR is shown at right.

cytogenetic CGH, which only has a resolution of ~10 Mb (11). SNP mapping arrays are a relatively new technique with the capacity to characterize both copy number and allelic imbalance aberrations in a single assay (12). In this study, we use the Affymetrix GeneChip Mapping 500K SNP arrays, which have a median intermarker distance of 2.5 kb, to analyze 31 primary epithelial ovarian tumors at high resolution.

Materials and Methods

Clinical samples. Ovarian cancer biopsies from three of the most common epithelial subtypes (11 serous, 9 mucinous, and 11 endometrioid) were obtained from women undergoing surgery for primary ovarian cancer at hospitals in the south of England (Supplementary Table S1). Tumor DNA was extracted from fresh-frozen tissue. To enhance copy number detection and improve genotype calling, all tumors were manually needle microdissected to ensure they contained >85% epithelial tumor cells. To visualize the epithelial tumor component, 10-μm sections were stained with cresyl violet (0.2%) and DNA was extracted using the DNeasy kit (Qiagen) following the manufacturer's instructions (13). To reduce the potential for admixing different clonal outgrowths, tumor cells were extracted from as few consecutive sections as possible (5–15) and none of the cancer cells were separated by >11 mm. Normal DNA was extracted from matching peripheral blood samples as described previously (14). Appropriate institutional ethics committees approved the collection and use of tissues for this study and ethics approval for this project was obtained from the Peter MacCallum Cancer Centre.

Mapping 500K arrays and data analysis. Two 250K Affymetrix SNP Mapping arrays (*Nsp*I and *Sty*I, Affymetrix) were used to analyze each tumor and matched normal sample using 250 ng of DNA per array as per the manufacturer's instructions and as described previously (15).

Data were first analyzed in GTYPE v4.0 (Affymetrix) to generate genotype information. The mapping algorithm call threshold setting was set at 0.26. "No call" genotypes were removed from allelic imbalance analysis; however, they were included in copy number determination because regions of homozygous deletion often contain no call genotypes. Copy number information was derived after first normalizing each tumor to its matched normal in CNAG v2.0 (16). For each sample, data from the 250K *Sty*I and 250K *Nsp*I arrays were combined and exported as text and Integrated Genome Browser (Affymetrix) compatible files. National Center for Biotechnology Information May 2004 Genome Build 35 was used for genome locations. Regions of copy number gain and loss were visually identified in Integrated Genome Browser and confirmed by interrogation of the raw copy number $\log_2$ ratios from CNAG v2 (16). Additional information on data analysis methods is provided in the Supplementary Data.

Statistical analyses were done using the Mann-Whitney test and the Spearman nonparametric correlation using GraphPad InStat (GraphPad Software). Fisher's exact test was carried out in R. $P < 0.05$ was considered significant.

Quantitative real-time PCR. PCR was carried out incorporating SYBR Green dye (ABgene UK) using the RotorGene 3000 (Corbett Research Australia). A standard curve of four normal DNA or human ovarian surface epithelium cell line cDNA dilutions was amplified in duplicate for each primer set and experiment. All samples were amplified at least twice. Because of the uncertainties involved when selecting a single locus primer set for genomic quantitative genomic real-time PCR (QPCR) that will have a normal copy number in every sample, copy number ratios were determined by comparison of the test primer with a control primer set for Alu repeat sequences (17) using the formula $2^{|(\text{CTAlu sample} - \text{CTTest sample}) - (\text{CTAlu standard} - \text{CTTest standard})|}$, where CT is the cycle number at which the amplification curve crosses the threshold. Expression QPCR used two housekeeping genes, *PGK1* and *HPRT1*, as control primers and the human ovarian surface epithelium cell line standard curve was used to calculate the expression ratio of fibroblast growth factor 1 (*FGFR1*) to the control genes in the tumor cDNA. Primer sequences are available on request.

Results and Discussion

500K array performance. Excellent call rates were obtained from all samples with an average of 91.1% for the tumor and

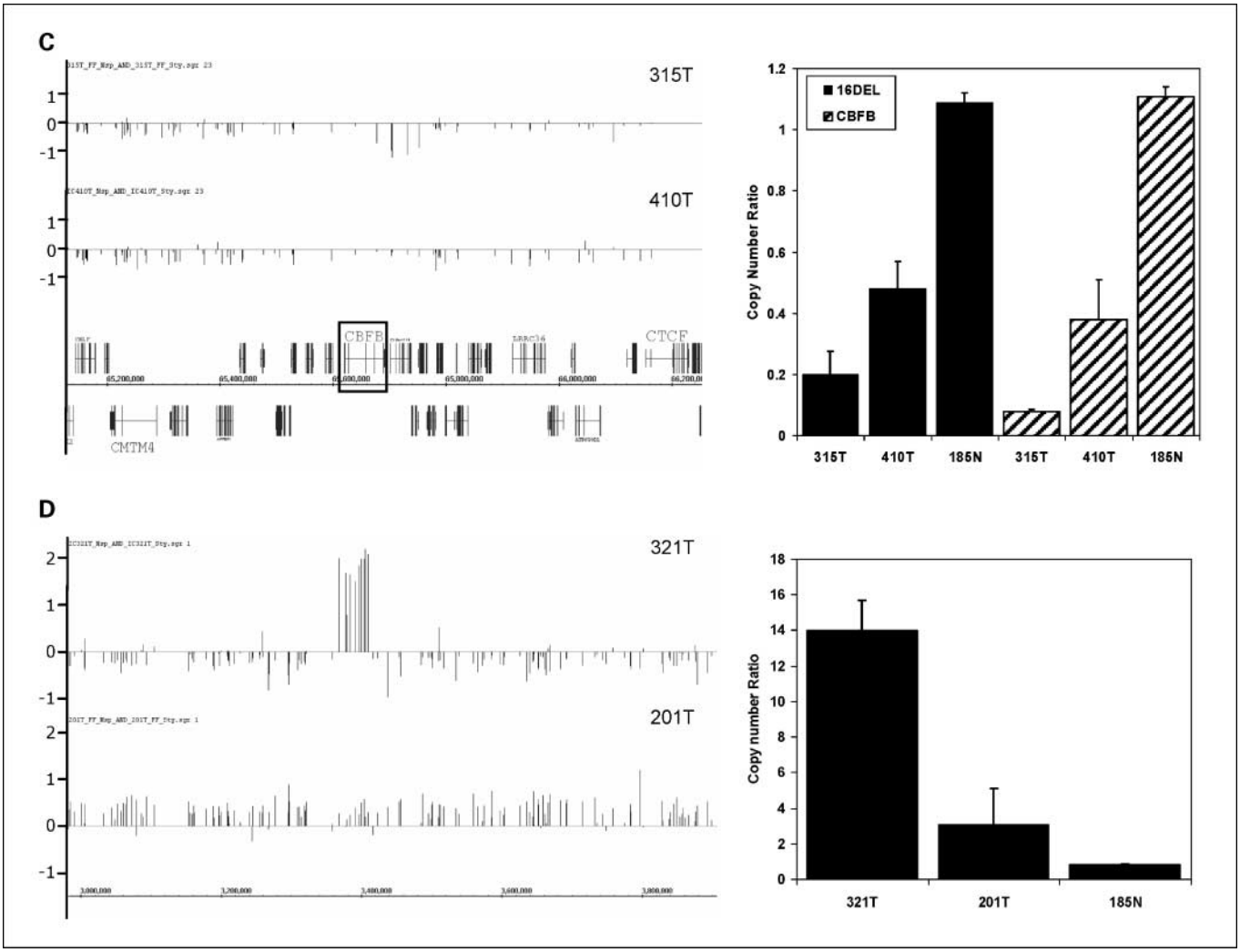

Fig. 1. *Continued. C,* chr16. Left, SNP $\log_2$ values of sample 315T, which shows a homozygous deletion on chromosome 16, and 410T, which shows hemizygous loss. At right is shown QPCR results using two sets of primers to the deletion on chromosome 16, confirming homozygous loss (315T), hemizygous loss (410T), and no change (185N). *D,* chr18. Left, SNP values of 321T (amplification) and 201T (low-level gain). At right are the confirmatory QPCR results. Columns, mean; bars, SE.

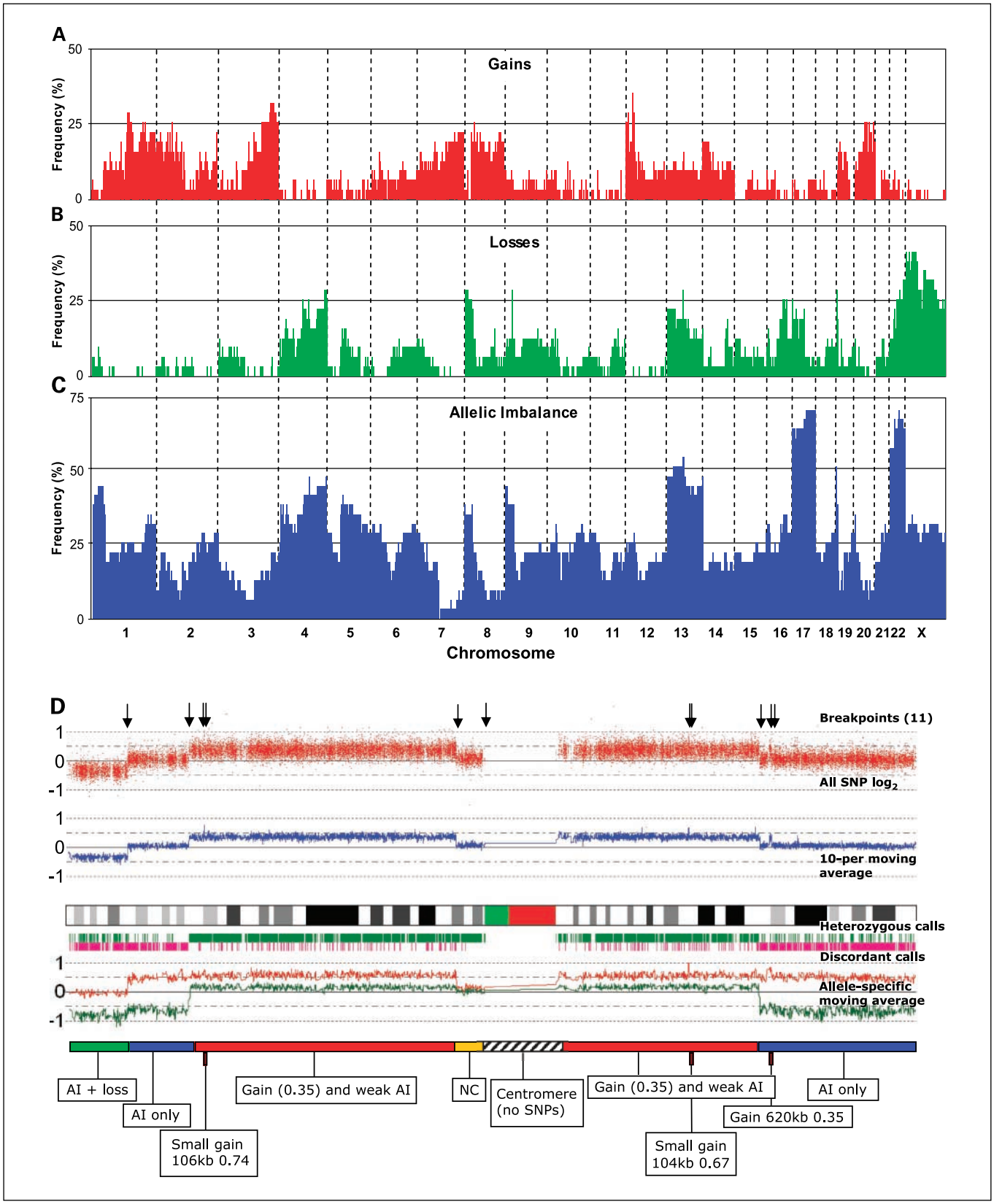

Fig. 2. Frequency of copy number changes and loss of heterozygosity in 31 ovarian cancer samples. *A,* gains. *B,* losses. *C,* allelic imbalance. Data are organized by chromosome cytoband from p terminus to q terminus. Any gain or loss within a cytoband is scored. *D,* CNAG output of 281T chromosome 1 to show the complexity of copy number and allelic imbalance alterations. Red dots, individual SNP $\log_2$ ratios; arrows, chromosome copy number breakpoints; blue line, 10-SNP moving average; green bars, tumor heterozygous genotypes; pink bars, paired tumor-normal discordant genotypes; red and green lines, allele-specific $\log_2$ intensity ratios. AI, allelic imbalance; NC, no change.

92.3% for the matching normal DNA. Copy number changes were clearly evident, with some samples showing just a few relatively small interstitial changes (Supplementary Fig. S1A), whereas others harbored numerous complex alterations on every chromosome (Supplementary Fig. S1B).

Criteria for assessing genome-wide copy number changes and allelic imbalance. Several variables may compromise the accuracy of copy number estimations, including germ-line copy number variation (18–20), germ-line polymorphisms that interfere with the restriction digestion, and random noise. To counter the first two sources of potential error, we analyzed matching normal DNA for all cases and normalized all tumors against their matching normal DNA. The data point for each SNP was a mean-normalized hybridization intensity value represented as a $\log_2$ ratio relative to normal. This strategy should effectively mask any germ-line copy number variations and emphasize only somatically acquired copy number changes. To compensate for random noise, a specific SNP was considered as showing gain or loss if the average $\log_2$ ratio of the region containing that SNP was at least ± 0.3.

As it is presently unclear how reliable individual SNP copy number values are, we designed QPCR primers to small regions of copy number change to assess if they were true copy number alterations. Four small regions of copy number change ranging from 59 to 350 kb and containing 5 to 45 SNPs were all validated (Fig. 1). We have taken a conservative stance and designated a region as showing copy number gain or loss if it was represented by at least five consecutive SNPs with average copy number values of ± 0.3 ($\log_2$) and encompassed regions of >25 kb in size. Although this is an arbitrary filter, a recent array CGH study of neuroblastoma was able to detect copy number alterations down to 25 kb with 390,000 oligonucleotide probes (21), suggesting that a 25-kb cutoff would exclude most false-positive copy number changes. We recognize that this filter may have removed some genuine very small copy number changes and that not all the >25-kb regions are necessarily accurate.

In addition to using the data to identify cancer gene loci, we also wished to gauge the level of genetic instability using the number of breakpoints as a surrogate measure. Because of sample polyploidy and/or tumor heterogeneity, there were many regions that did not exceed the threshold set for gains and losses ($\log_2$ ratio of ± 0.3) and yet a copy number breakpoint could readily be discerned, for example, a shift in average $\log_2$ copy number ratio from -0.2 to 0.2 (Supplementary Fig. S2A). Therefore, we have defined copy number breakpoints as any significant shift in average copy number ratio between adjacent regions regardless of whether the average ratio for either region exceeded the $\log_2$ copy number ratio ± 0.3 threshold (see Supplementary Methods for more information). For each sample, the "breakpoint index" was the sum of all copy number breakpoints observed.

Allelic imbalance was detected by analysis of the allele-specific data in CNAG v2.0. This analysis could readily determine large regions of allelic imbalance but had limited capacity for detection of small regions when unaccompanied by copy number change.

Allelic imbalance and copy number changes organized by cytoband. A genome-wide summary of the copy number changes and allelic imbalance for all samples, organized by cytoband, is shown in Fig. 2A to C and Supplementary Table S2. Our data are in good agreement with the most frequently reported alterations, such as gain at 3q26 (35%), loss at 4q (29%), and allelic imbalance at chromosomes 13 (55%) and 17 (71%; ref. 22). However, we identified many more copy number changes than previous studies. For example, losses at 8p23-22 (29%), 9p21.3 (29%), 19p13 (29%), 22q (32%), and Xp (42%) were all detected at higher frequency compared with the online CGH and cytogenetic database Progenetix.net (19%, 12%, 7%, 15%, and 14%, respectively).[4] Gains at 12p12.1, which encompass *KRAS*, were also detected frequently in our sample set, with 35% of tumors showing gains compared with 18% in the Progenetix.net samples. Of the 528 cases in the Progenetix database, 75% were analyzed by cytogenetic CGH, which has an approximate resolution of 10 Mb. Our higher detection rate is therefore most likely due to the enhanced resolution of the Mapping 500K arrays. For example, of the 11 tumors showing copy number gains at 12p12.1 (containing *KRAS*), 4 encompassed regions of <5 Mb that potentially would have been missed using other techniques. Similarly, losses at 9p21.3 (containing *CDKN2A*) were detected in five samples and ranged in size from 370 kb to 4.2 Mb.

It is interesting to note that, although allelic imbalance often coincided with copy number loss in our samples, this was not always the case, as 35% of all allelic imbalance regions showed no change in copy number. For example, allelic imbalance was concomitant with copy number loss at 13q14.13-14.2 and 22q13.1-13.31 in 19% and 26% of tumors, respectively, whereas allelic imbalance alone was observed across the same regions in 35% and 42% of tumors. Thus, copy number neutral allelic imbalance that occurs through chromosome reduplication or mitotic recombination is a relatively frequent event in ovarian cancers. Almost all allelic imbalance breakpoints (95%) coincided with copy number breakpoints in our data set. However, 18 regions of allelic imbalance had boundaries that were independent of any copy number breakpoint, suggesting a mechanism of mitotic recombination for these

Table 1. Distribution of breakpoint indices, comparing low breakpoint index with intermediate/high breakpoint index

	Low breakpoint index	Intermediate/High breakpoint index	*P*
Grade 1	3	0	
Grade 2	5	7	
Grade 3	3	13	0.02
Stage I	9	2	
Stage II	1	2	
Stage III and IV*	1	12	0.0005
Serous	0	11	
Endometrioid	4	7	
Mucinous	7	2	0.0007

NOTE: *P* values were calculated using the Fisher's exact test.
*Stages III and IV were combined because there was only one stage IV tumor.

[4] http://www.progenetix.de

Table 2. Small regions of gain and loss in ovarian tumors

	No. regions	Mean size (bp)	Size range (kb)	Median no. genes	No. regions with no genes	No. regions with known cancer genes
Low-level gain (<0.5 Mb)	200	271,639	26-496	1	43	11
High-level gain (<0.5 Mb)	33	207,685	59-484	1	13	1
Hemizygous loss (<0.5 Mb)	127	209,032	31-495	1	52	5
Homozygous loss (<0.5 Mb)	27	151,389	25-421	1	10	5

NOTE: Log_2 copy number ratios: (*a*) low-level gain, 0.3 to 0.99; (*b*) high-level gain, ≥1.0; (*c*) hemizygous loss, -0.3 to -0.69; (*d*) homozygous loss, ≤-0.7. Known cancer genes from Cancer Gene Census (http://www.sanger.ac.uk/genetics/CGP/Census/).

aberrations. For example, allelic imbalance of the distal part of chromosome 9p in five mucinous tumors all terminated in copy number neutral regions. The high resolution of copy number detection coupled with the ability to simultaneously identify allelic imbalance uncovered some highly complex structural and copy number alterations. For example, tumor 281 shows 11 breakpoints across chromosome 1 with large regions containing allelic imbalance plus copy number loss (due to loss of one allele), allelic imbalance without copy number change (due to loss of one allele and reduplication of the other), and weak allelic imbalance with low level copy number gain (due to duplication of one allele). Additionally, chromosome 1 contained three small (106, 104, and 630 kb) regions of copy number gain, one of which was located in a region showing allelic imbalance (Fig. 2D).

Copy number breakpoints are frequent in serous and endometrioid ovarian cancers. The breakpoint index was counted for each tumor as a surrogate measure of genomic instability (Supplementary Fig. S2B). The breakpoint index ranged from 0 to 818 with a median of 95. This frequency is high compared with most other CGH studies of ovarian cancers that have reported an average of 10 to 20 breakpoints (copy number changes of 5-10 per tumor), with the most complex case reported showing at most 62 breakpoints (arising from 31 copy number alterations; ref. 23). In our data set, individual tumors showed great variation in the number of changes (compare Supplementary Fig. S1A and B) and could be broadly classified into three categories: those with a low breakpoint index (<15 chromosome breakpoints, 11 samples), those with an intermediate breakpoint index (33-305 breakpoints, 18 samples), and rare samples that showed high numbers of copy number alterations (>600 breakpoints, 2 samples). The proportion of tumors in the low breakpoint index versus intermediate/high breakpoint index varied between different histologic subtypes (Table 1). The samples with the fewest alterations were predominantly cancers of the mucinous histologic subtype. Endometrioid tumors seemed to fall in the low-intermediate range, whereas serous-type malignancies always showed several alterations and were the only tumor type represented in the high-level breakpoint group. This pattern is consistent with studies that have shown more CGH alterations in serous-type tumors compared with nonserous tumors (2, 24). There was an increasing number of breakpoints with advanced-stage and higher-grade tumors consistent with a large CGH study of ovarian cancer (4).

Ovarian cancers harbor a surprisingly high number of small regions of copy number gain and loss. The most striking feature of the data is the prevalence of small (arbitrarily defined as regions of <500 kb) copy number changes (Table 2; Supplementary Table S3): 387 small regions of copy number gain or loss were identified among the 31 tumors with the most common alteration being low-level gain (average log_2 copy number ratio between 0.3 and 1) followed by hemizygous loss (average log_2 copy number ratio between -0.3 and -0.7). The number of small regions of copy number change (Fig. 3) as well as the relative proportion of gains to losses varied widely between tumors. There were four groups of small copy number change: (*a*) those tumors with <10 changes (17 tumors), (*b*) those with changes equally distributed between gains and losses (4 tumors), (*c*) those with more losses than gains (3 tumors), and (*d*) those with more gains than losses (7 tumors).

The number of small regions of copy number gains and loss correlated with the number of larger-scale changes ($P = 0.002$; Spearman $r = 0.54$). These larger alterations were counted as copy number increase or decrease comprising ≥50% of a chromosome arm, including discontinuous changes (e.g.,

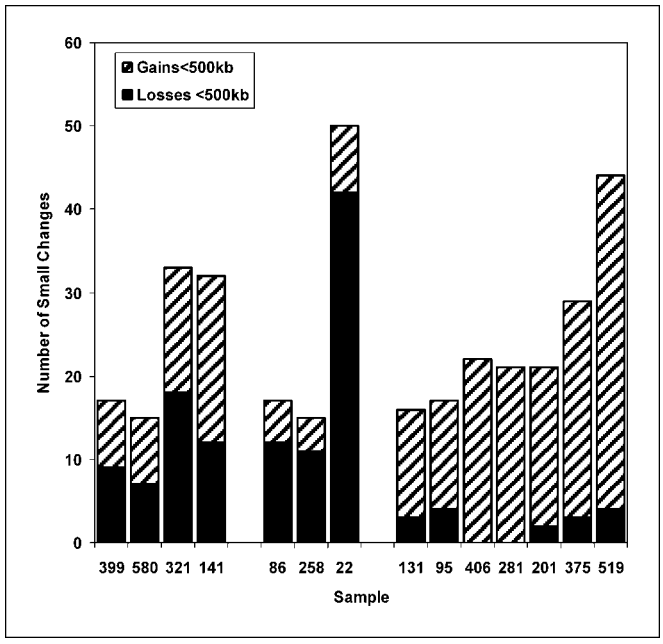

Fig. 3. Distribution of small regions of copy number change. Only tumors with >10 small (<500 kb) regions of copy number change are shown. Tumors are grouped into those with even numbers of gains and losses, more losses than gains, and more gains than losses.

Table 3. Small regions (<500 kb) of gain and loss containing known cancer genes

Sample	Chromosome	Cytoband	Start (Mb)	Size (kb)	Mean SNP $\log_2$ value	Gene list
High-level gain						
258	8	p12-p11.23	38.38	350	1.50	***FGFR1***, *FLJ43582*
Low-level gain						
519	1	p36.32	2.765	297	0.44	***PRDM16***, *ACTRT2*
131	2	q13-q14.1	113.65	200	0.34	***PAX8***, *PSD4*
281	8	p12	38.010	244	0.41	***WHSC1L1***, *ASH2L, BAG4, DDHD2, STAR, LSM1, PPAPDC1B, EIF4EBP1*
519	8	p12	30.86	380	0.42	***WRN***, *PURG*
201	9	q34.13	131.055	446	0.30	***NUP214***, *PPAPDC3, POMT1, UCK1, FAM78A, RAPGEF1*
141	12	p12.1	25.13	395	0.56	***KRAS***, *CASC1, LRMP, LYRM5, IFLTD1*
519	14	q23.3	66.147	175	0.34	***GPHN***
406	17	p13.1	9.736	99	0.74	***GAS7***, *RCV1*
406	17	p13.1	9.835	200	0.39	***GAS7***
375	19	q12	34.55	470	0.40	***CCNE1***, *POP4, PLEKHF1, C19orf12*
410	19	p13.12-p13.11	15.925	343	0.43	***TPM4***, *RAB8A, HSH2D, FAM32A, AP1M1, CIB3, OR10H4, FLJ25328, LOC126536*
Homozygous loss						
86	13	q14.1	47.817	141	-1.12	***RB1***
258	13	q14.1	47.73	109	-1.10	***RB1***
315	13	q14.1	47.854	216	-1.00	***RB1***, *P2RY5, RCBTB2*
315	16	q22.1	65.64	138	-1.10	***CBFB***, *LIN10, FBXL8, HSF4, NOL3, TRADD, MGC4655, LOC283849*
141	17	p12	11.856	344	-0.90	***MAP2K4***
Hemizygous loss						
258	2	p16.3	47.825	445	-0.40	***MSH6***, *FBXO11*
519	6	q21	108.78	356	-0.49	***FOXO3A***, *LACE1*
40	9	p21.3	21.778	416	-0.65	***CDKN2A***, *CDKN2B, MTAP*
375	10	p12.31	21.948	377	-0.37	***MLLT10***, *DNAJC*
258	X	q13.1	69.94	440	-0.54	**MLLT7**, *MED12, NLGN3, GJB1,* **NONO**, *ITGB1BP2, TAF1, SNX12, LOC15883, IL2RG, ZMYM3*

NOTE: Cancer genes from Cancer Gene Census are shown in bold italicized type. Size is the maximum possible size. Genome locations from National Center for Biotechnology Information Build 35.

whole chromosome arm loss interrupted by a small return to normality). Similarly, large-scale allelic imbalance correlated with small copy number changes (P = 0.0035; Spearman r = 0.50). Consequently, we speculate that the large-scale and small-scale copy number and allelic imbalance alterations may arise through a common pathway in many tumors. However, this trend was not always observed; for example, tumor 594 seemed to be hypodiploid with a chromosome modal number of 32 (14 large-scale copy number changes) but had only four breakpoints, contrasting with tumor 406, which had 181 breakpoints but only one large-scale chromosome copy number change. These tumors may have alternate chromosome instability mechanisms for the disparate occurrence of large-scale versus small-scale alterations. Clearly, gain or loss of entire chromosomes may arise through failure of mitotic segregation checkpoints (numerical chromosome instability) and this type of instability may be present in tumors such as 594. However, purely structural chromosome instability could arise through defective recombination and repair pathways, although it is unclear how differences in the pathways affected would give rise to some samples showing more gains than losses and vice versa. A mechanism for combined structural and numerical chromosome instability has been described via telomere aggregation (25), leading to both breakage-fusion-bridge cycles (=structural rearrangements) and also missegregation and micronucleus formation (=aneuploidy). One of the causes of this phenotype was overexpression of MYC, and interestingly, the two samples with high-level amplification of the *MYC* gene, 375 and 141, had moderately high breakpoint levels, both small-scale and large-scale structural aberrations, and were also estimated to be aneuploid.

Small regions of copy number gain and loss often encompass genes implicated in cancer. Most of the small regions of copy number gain and loss identified contained at least one known gene and 22 regions contained known cancer genes [as defined by Futreal et al. (26); Table 3]. This represents a significant enrichment of cancer genes within the small regions of copy number gain or loss (P = 0.003). Three samples had small regions of deletion affecting the retinoblastoma 1 (*RB1*) locus. *RB1* has previously been shown to be inactivated in a subset of ovarian tumors (27). The maximum sizes of the deletions in these three tumors are 109, 130, and 216 kb (Fig. 1A) and are clearly homozygous, showing average SNP $\log_2$ ratios of -1.0 to -1.2. We undertook QPCR to verify these deletions. As the three deletions were not overlapping, two sets of primers were designed, one to the 5′ end (*RB1* intron 1) and one to the 3′ end (*RB1* exon 27) of *RB1*. The copy number ratios obtained by QPCR were <0.1 compared with normal, confirming that all three tumors harbored homozygous deletions in *RB1* (Fig. 1A). For comparison, tumor 201 shows hemizygous loss at *RB1* (average SNP $\log_2$ ratio of -0.4, average QPCR copy number ratio of 0.45).

A 350-kb high-level amplification was identified in the endometrioid tumor 258 on 8p (Fig. 1B). Only two genes are annotated within the amplification: an uncharacterized cDNA and *FGFR1*. QPCR confirmed the amplification and estimated the copy number ratio to be 10 (Fig. 1B). Expression of *FGFR1* was also increased in sample 258. *FGFR1* is a known oncogene shown to be amplified in a variety of cancers, including a small proportion of ovarian tumors (28, 29). *FGFR1* also showed copy number gain in three other endometrioid tumors and two serous tumors. The specificity of the 8p amplification in tumor 258 suggests that *FGFR1* is the primary target in ovarian cancer, and thus, this amplicon may be distinct from the complex 8p amplifications observed in 15% of breast cancers that may have several other targets (30).

Another small region of copy number loss was located on chromosome 16q22.1 in tumor 315 (Fig. 1C). This deletion was confirmed by two sets of QPCR primers designed to the region, one set for each end (Fig. 1C). Comprising only five SNPs, the homozygous deletion nevertheless covered a maximum size of 137 kb and eight genes. One of these genes is *core-binding factor β* (*CBFB*), a transcriptional activator with a known function in acute myelogenous leukemia (31); however, its role in ovarian cancer is unknown. In acute myelogenous leukemia, *CBFB* is mutated by deletion or chromosome rearrangement, leading to dysregulation of its ability to activate transcription with its binding partner RUNX1. Although seven tumors in our data set showed hemizygous loss at 16q22.1, no other homozygous deletions or chromosome breakpoints targeted this gene. Allelic imbalance at 16q22.1 was found in 30% of our tumors (excluding the tumor with the homozygous deletion, 315). Whether the homozygous deletion is indicative of a wider role for *CBFB* in ovarian cancer through point mutation or methylation of the remaining allele in the tumors with allelic imbalance remains to be determined.

As well as genes with a known role in ovarian cancer, such as *KRAS* and *RB1*, Table 3 lists several cancer genes, such as *CBFB*, whose role in ovarian cancer is plausible but not yet established. Of particular interest is *MutS homologue 6* (*MSH6*), which is one of the genes responsible for the hereditary nonpolyposis colorectal cancer familial cancer syndrome that includes endometrioid ovarian cancer as part of the disease spectrum (32). Tumor 258, which is of the endometrioid histologic subtype, harbored a small region of hemizygous loss across *MSH6*, although full sequencing of the coding region in this tumor did not reveal a second inactivating mutation and there was no evidence of promoter hypermethylation (data not shown), raising the possibility that this gene could be affected by haploinsufficiency. Alternatively, the other gene in the region, *F-box protein 11* (*FBXO11*), could be the target of this deletion. Another possibility is that this deletion is not pathogenic and may be a passenger alteration arising through genetic instability in the tumor.

A small deletion in the mucinous tumor 40 targeted *CDKN2A* (*p16*) on chromosome 9p. This gene was targeted most frequently in mucinous tumors. Four of the five tumors that had deletions under 5 Mb in size at the *CDKN2A* locus were mucinous. Overall, 6 of 9 (67%) of mucinous tumors had allelic imbalance of 9p and/or specific deletions of the *CDKN2A* region compared with only 6 of 22 (27%) of endometrioid and serous tumors ($P = 0.056$, Fisher's exact test). Previous studies of *CDKN2A* in ovarian cancer have found mutations or deletions predominantly in mucinous tumors, but the prevalence may have been underestimated in this subtype, as few studies have investigated ovarian mucinous tumors specifically (33, 34). An allelic imbalance study of ovarian mucinous tumors found high levels of allelic imbalance at 9p (80%; ref. 35). The mucinous tumors studied here were generally different from the serous and endometrioid tumors. Not only did they have the fewest copy number alterations, but they also had a different pattern of loss of heterozygosity (Supplementary Table S4). Only the mucinous tumors with the highest breakpoint levels, 403 and 321, had loss of heterozygosity at all three of the most frequently affected regions in serous and endometrioid tumors: chr22q, chr17, and chr13q (20 of 22, 18 of 22, and 16 of 22 of the nonmucinous tumors, respectively). This may indicate that these alterations only affect some high-grade ovarian mucinous tumors (403 is grade 3 and 321 is grade 2) or they may represent a different subtype of mucinous tumor. Few copy number differences were identified between serous and endometrioid tumors. The biggest difference was an increase in chr1 gain in 5 of 11 (55%) serous over 2 of 11 (18%) endometrioid; however, this was not significant due to the small number of samples ($P = 0.18$, Fisher's exact test).

A 344-kb homozygous deletion specifically targeting the *mitogen-activated protein kinase kinase 4* (*MAP2K4*) gene was detected in one tumor (tumor 141). A larger, 1.79-Mb homozygous deletion including this gene was also observed in a second tumor (tumor 22). Somatic mutations in *MAP2K4*, located on 17p12, have been identified in a low percentage of several different tumor types, including breast, colon, and pancreatic (36). No point mutations in *MAP2K4* in ovarian cancer have been found; however, allelic imbalance of 17p is an extremely frequent event (65% in our data set) and may not always target *TP53* alone (located at 17p13) particularly because the frequency of *TP53* mutation in ovarian cancer is ~48% (IARC *TP53* mutation database, R11; ref. 37). A recent study identified deletions of *MAP2K4* in two serous tumors and reduced expression in 96 of 128 serous cancers (38). Both tumors with homozygous deletions in our data set were of the serous type. *MAP2K4* has also been hypothesized to be a suppressor of metastasis in ovarian and prostate cancer (39).

Many (118) small regions contained no known genes (based on Integrated Genome Browser RefSeq genes, assembly hg17, May 2004), although some of these did have uncharacterized cDNAs or predicted transcripts by Genscan. None of these regions held known microRNAs; however, two small regions of loss and seven regions of low-level gain contained microRNAs as well as 1 to 23 known genes. microRNAs are predicted to act as tumor suppressor genes and oncogenes (40) so that copy number alterations that affect microRNA gene expression might affect the transcriptional regulation of their target genes.

Conclusion

This study is, to our knowledge, the most extensive and highest-resolution study of genetic alterations in primary ovarian cancer to date. Our data confirm the cytogenetically complex and heterogeneous nature of ovarian cancer but also reveal a surprising number of small regions of copy number change. Some of these regions contain cancer genes known to

be involved in ovarian cancer and others hold genes that have a known role in other cancer types but which have yet to be established as ovarian cancer genes. Many of the small regions of copy number change lie within genomic intervals frequently affected in ovarian cancer but for which the targeted gene has yet to be identified. Therefore, our characterization of 31 ovarian cancer genomes at such high resolution has yielded several novel candidate genes for future analyses and reveals a more complex genomic structure of ovarian tumors than previously determined.

References

1. Hoglund M, Gisselsson D, Hansen GB, Sall T, Mitelman F. Ovarian carcinoma develops through multiple modes of chromosomal evolution. Cancer Res 2003;63:3378–85.
2. Hauptmann S, Denkert C, Koch I, et al. Genetic alterations in epithelial ovarian tumors analyzed by comparative genomic hybridization. Hum Pathol 2002;33:632–41.
3. Iwabuchi H, Sakamoto M, Sakunaga H, et al. Genetic analysis of benign, low-grade, and high-grade ovarian tumors. Cancer Res 1995;55:6172–80.
4. Kiechle M, Jacobsen A, Schwarz-Boeger U, Hedderich J, Pfisterer J, Arnold N. Comparative genomic hybridization detects genetic imbalances in primary ovarian carcinomas as correlated with grade of differentiation. Cancer 2001;91:534–40.
5. Sonoda G, Palazzo J, du Manoir S, et al. Comparative genomic hybridization detects frequent overrepresentation of chromosomal material from 3q26, 8q24, and 20q13 in human ovarian carcinomas. Genes Chromosomes Cancer 1997;20:320–8.
6. Gaasenbeek M, Howarth K, Rowan AJ, et al. Combined array-comparative genomic hybridization and single-nucleotide polymorphism-loss of heterozygosity analysis reveals complex changes and multiple forms of chromosomal instability in colorectal cancers. Cancer Res 2006;66:3471–9.
7. Shadeo A, Lam WL. Comprehensive copy number profiles of breast cancer cell model genomes. Breast Cancer Res 2006;8:R9.
8. Janne PA, Li C, Zhao X, et al. High-resolution single-nucleotide polymorphism array and clustering analysis of loss of heterozygosity in human lung cancer cell lines. Oncogene 2004;23:2716–26.
9. Lambros MB, Fiegler H, Jones A, et al. Analysis of ovarian cancer cell lines using array-based comparative genomic hybridization. J Pathol 2005;205:29–40.
10. Kim SW, Kim JW, Kim YT, et al. Analysis of chromosomal changes in serous ovarian carcinoma using high-resolution array comparative genomic hybridization: potential predictive markers of chemoresistant disease. Genes Chromosomes Cancer 2007;46:1–9.
11. Bentz M, Plesch A, Stilgenbauer S, Dohner H, Lichter P. Minimal sizes of deletions detected by comparative genomic hybridization. Genes Chromosomes Cancer 1998;21:172–5.
12. Huang J, Wei W, Zhang J, et al. Whole genome DNA copy number changes identified by high density oligonucleotide arrays. Hum Genomics 2004;1:287–99.
13. Jiang X, Hitchcock A, Bryan EJ, et al. Microsatellite analysis of endometriosis reveals loss of heterozygosity at candidate ovarian tumor suppressor gene loci. Cancer Res 1996;56:3534–9.
14. Mullenbach R, Lagoda PJ, Welter C. An efficient salt-chloroform extraction of DNA from blood and tissues. Trends Genet 1989;5:391.
15. Jacobs S, Thompson ER, Nannya Y, et al. Genome-wide, high-resolution detection of copy number, loss of heterozygosity, and genotypes from formalin-fixed, paraffin-embedded tumor tissue using microarrays. Cancer Res 2007;67:2544–51.
16. Nannya Y, Sanada M, Nakazaki K, et al. A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res 2005;65:6071–9.
17. Nicklas JA, Buel E. Development of an Alu-based, real-time PCR method for quantitation of human DNA in forensic samples. J Forensic Sci 2003;48:936–44.
18. Iafrate AJ, Feuk L, Rivera MN, et al. Detection of large-scale variation in the human genome. Nat Genet 2004;36:949–51.
19. Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. Science 2004;305:525–8.
20. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature 2006;444:444–54.
21. Selzer RR, Richmond TA, Pofahl NJ, et al. Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. Genes Chromosomes Cancer 2005;44:305–19.
22. Baudis M, Cleary ML. Progenetix.net: an online repository for molecular cytogenetic aberration data. Bioinformatics 2001;17:1228–9.
23. Hu J, Khanna V, Jones MW, Surti U. Comparative study of primary and recurrent ovarian serous carcinomas: comparative genomic hybridization analysis with a potential application for prognosis. Gynecol Oncol 2003;89:369–75.
24. Osterberg L, Levan K, Partheen K, Helou K, Horvath G. Cytogenetic analysis of carboplatin resistance in early-stage epithelial ovarian carcinoma. Cancer Genet Cytogenet 2005;163:144–50.
25. Louis SF, Vermolen BJ, Garini Y, et al. c-Myc induces chromosomal rearrangements through telomere and chromosome remodeling in the interphase nucleus. Proc Natl Acad Sci U S A 2005;102:9613–8.
26. Futreal PA, Coin L, Marshall M, et al. A census of human cancer genes. Nat Rev Cancer 2004;4:177–83.
27. Liu Y, Heyman M, Wang Y, et al. Molecular analysis of the retinoblastoma gene in primary ovarian cancer cells. Int J Cancer 1994;58:663–7.
28. Theillet C, Adelaide J, Louason G, et al. FGFRI and PLAT genes and DNA amplification at 8p12 in breast and ovarian cancers. Genes Chromosomes Cancer 1993;7:219–26.
29. Zhao X, Weir BA, LaFramboise T, et al. Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. Cancer Res 2005;65:5561–70.
30. Gelsi-Boyer V, Orsetti B, Cervera N, et al. Comprehensive profiling of 8p11-12 amplification in breast cancer. Mol Cancer Res 2005;3:655–67.
31. Claxton DF, Liu P, Hsu HB, et al. Detection of fusion transcripts generated by the inversion 16 chromosome in acute myelogenous leukemia. Blood 1994;83:1750–6.
32. Miyaki M, Konishi M, Tanaka K, et al. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. Nat Genet 1997;17:271–2.
33. Campbell IG, Foulkes WD, Beynon G, Davis M, Englefield P. LOH and mutation analysis of CDKN2 in primary human ovarian cancers. Int J Cancer 1995;63:222–5.
34. Milde-Langosch K, Ocon E, Becker G, Loning T. p16/MTS1 inactivation in ovarian carcinomas: high frequency of reduced protein expression associated with hyper-methylation or mutation in endometrioid and mucinous tumors. Int J Cancer 1998;79:61–5.
35. Feltmate CM, Lee KR, Johnson M, et al. Whole-genome allelotyping identified distinct loss-of-heterozygosity patterns in mucinous ovarian and appendiceal carcinomas. Clin Cancer Res 2005;11:7651–7.
36. Teng DH, Perry WL III, Hogan JK, et al. Human mitogen-activated protein kinase kinase 4 as a candidate tumor suppressor. Cancer Res 1997;57:4177–82.
37. Olivier M, Eeles R, Hollstein M, Khan MA, Harris CC, Hainaut P. The IARC TP53 database: new online mutation analysis and recommendations to users. Hum Mutat 2002;19:607–14.
38. Nakayama K, Nakayama N, Davidson B, et al. Homozygous deletion of MKK4 in ovarian serous carcinoma. Cancer Biol Ther 2006;5:630–4.
39. Yamada SD, Hickson JA, Hrobowski Y, et al. Mitogen-activated protein kinase kinase 4 (MKK4) acts as a metastasis suppressor gene in human ovarian carcinoma. Cancer Res 2002;62:6717–23.
40. Caldas C, Brenton JD. Sizing up miRNAs as cancer genes. Nat Med 2005;11:712–4.